THE BELL SYSTEM

Technical Journal

DEVOTED TO THE SCIENTIFIC AND ENGINEERING ASPECTS OF ELECTRICAL COMMUNICATION

VOLUME XXXI NOVEMBER 1952 NUMBER 6

THE BELL SYSTEM TECHNICAL JOURNAL

THE BELL SYSTEM TECHNICAL JOURNAL is published six times a year by the American Telephone and Telegraph Company, 195 Broadway, New York 7, N. Y. Cleo F. Craig, President; Carroll O. Bickelhaupt, Secretary; Donald R. Belcher, Treasurer. Subscriptions are accepted at $3.00 per year. Single copies are 75 cents each. The foreign postage is 65 cents per year or 11 cents per copy. Printed in U. S. A.

THE BELL SYSTEM
TECHNICAL JOURNAL

VOLUME XXXI NOVEMBER 1952 NUMBER 6

A New General Purpose Relay for Telephone Switching Systems

By ARTHUR C. KELLER

(Manuscript received July 14, 1952)

This paper describes a new general purpose electromagnetic relay for use in telephone switching systems. It is a wire spring relay known as the AF type relay and, with variations which provide slow release or marginal characteristics, it is known as the AG and AJ relay, respectively. Fig. 1 shows a typical AF type relay, Fig. 2 shows all of the parts of the relay and Fig. 3 is a drawing showing the relay assembly.

1. BACKGROUND

The general purpose relay is one of the most important components of telephone switching systems.[1] These relays constitute the most repetitive building block in switching equipment. Since several million are produced annually, low manufacturing cost is extremely desirable. Also of prime importance are low operating and maintenance costs. General purpose relays are, therefore, under constant observation and study by the telephone operating companies as the users, by the Western Electric Company as the manufacturer, and by Bell Telephone Laboratories as the designer. The AF wire spring relay and its variations are the result of such studies.

A general purpose relay for telephone switching systems must meet a large number of diverse requirements. It must be capable of being assembled with any one of a variety of magnet coils having a wide range

[1] S. P. Shackleton and H. W. Purcell, "Relays in the Bell System", *Bell System Tech. J.*, Jan., 1924, p. 1.

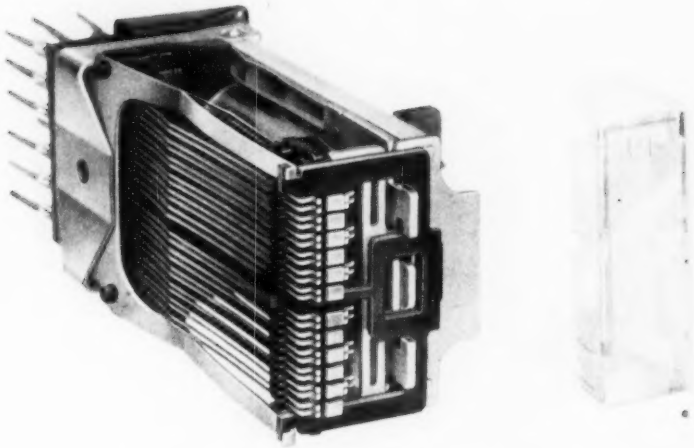


Fig. 1—AF type relay, with contact cover detached.

of resistance values and to operate contacts which vary from one pair to as many as fourteen or more. The basic relay design must also be capable of providing such features as fast operation and release, slow release, high sensitivity, heavy duty and marginal operation. These functions are performed satisfactorily in present crossbar switching systems by U, UA, UB and Y type relays.[2, 3, 4, 5] However, with an objective of a forty-year life for new switching systems and a trend toward unattended operation of switching offices, it is important to attain the best in the performance and reliability of relays.

The general purpose relay must be designed to produce the best economic balance, when used in telephone switching systems, so that the annual charges are minimized. The major ingredients of these annual charges are manufacturing expense, operating electrical power, speed of operation and release, space required and maintenance costs which include reliability and life.

2. REQUIREMENTS AND OBJECTIVES

The requirements for a new general purpose relay were initially broadly stated to be performance and maintenance at least equal to the

[2] H. N. Wagar, "The U-Type Relay", *Bell Lab. Record*, May, 1938, p. 300.
[3] H. M. Knapp, "The UB Relay", *Bell Lab. Record*, Oct., 1949, p. 355.
[4] F. A. Zupa, "The Y-Type Relay", *Bell Lab. Record*, May, 1938, p. 310.
[5] W. C. Slauson, "Improved U, UA and Y Type Relays", *Bell Lab. Record*, Oct., 1951, p. 466.

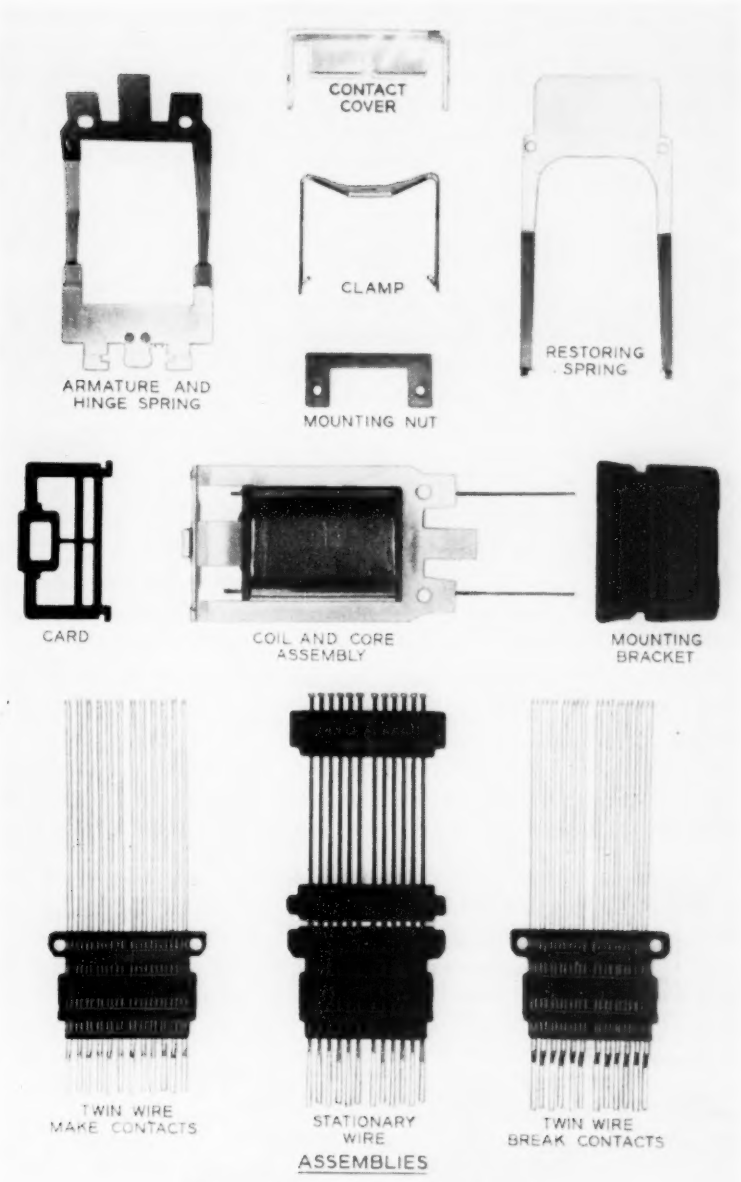


Fig. 2—Parts of the wire spring relay.

U and Y type relays but with substantially lower manufacturing costs. As the development of the relay proceeded, it became possible to expand the requirements without appreciably altering the expected relay cost. In particular, it became possible to design the new relay to operate and release faster or to use less electrical power, to operate more often before appreciable wear occurred, etc. The improved performance characteristics of the AF wire spring relay, as described later, are of equal economic importance to those associated with lower manufacturing cost.

The broad requirements were reduced to the following design objectives:

1. Lower cost—50 per cent of U type relay.
2. Reduced operating electrical power.
3. Faster operate and release times.
4. Long life—one billion operations.
5. Improved contact performance.

These broad design objectives do not specifically state a large number of other characteristics which must be at least as favorable as those of the U and Y relay family. This refers to such items as: space required, magnetic interference, wiring costs, contact combinations, field servicing and repairs.

3. DESIGN POSSIBILITIES

The studies of new relay design possibilities started with a careful review of the U type relay experience. In fact, much of the early thinking considered various modifications of U type and other existing relays. In general, these studies indicated that about half of the manufacturing cost of U type relays came from assembly and adjusting operations. Accordingly, these operations required major revision for a substantial

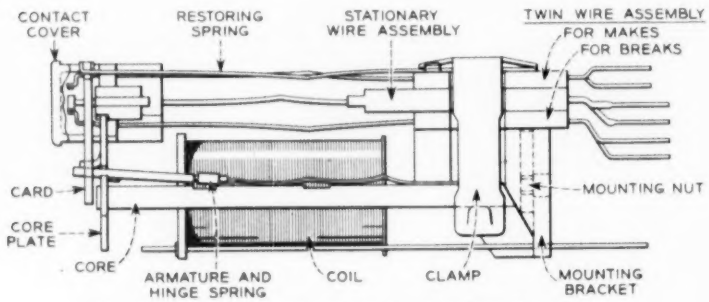


Fig. 3—Top view of the relay, showing location of parts.

cost reduction of the relay. It became evident that the development of new manufacturing methods as well as new designs were essential in reaching the ambitious objectives. For these reasons, the manufacturing engineers of the Western Electric Company were active participants in the development of the new relay from the beginning.

Many new forms of relay designs were considered and studied including such types as miniature, magnetic contact, piezoelectric, etc. As a result, one general form, first proposed by H. C. Harrison, gave the most promise of meeting the manifold requirements. This is the wire spring type characterized by the wire spring subassemblies with code card operation of pretensioned, low stiffness springs. Actually, the general form of the wire spring relay proposed by Mr. Harrison constitutes an entire new class of relays with many possible variations. These include various types of code card operation and various forms of contact operation, operated by any of a number of magnet structures.

The new class of relays has the following important advantages:

1. Pretensioned, low stiffness wire springs make possible (a) assembly to give close control of contact force without individual spring adjustment; and (b) essentially constant contact force throughout the life of the relay and its contacts.

2. Wire spring subassemblies make possible (a) favorable manufacture of a multiplicity of contact springs by molding; (b) lower assembly costs because fewer piece parts are needed; and (c) simple code card operation.

3. Code card operation makes possible (a) standardized and simple assembly; (b) accurate control of contact position; (c) essential elimination of locked contacts; (d) complete independence of twin contacts; and (e) simple means for providing a large number of contact combinations.

A continuous and comprehensive study was necessary of the characteristics and probable manufacturing costs of many forms of the wire spring relay family. As a result, after passing through several major designs, the basic design of the present relay was adopted. H. M. Knapp and C. F. Spahn proposed important features of this design. This form represented advantages over other types in

1. reducing the number and amount of dimensional variations controlling the contact gaps. In turn, this made possible smaller armature movement, shorter operating and release times and less chatter of the contacts;

2. reducing the number of code cards required to provide the large number of contact combinations needed in switching systems;

3. reducing the manufacturing and wiring costs;

4. increasing the mechanical life.

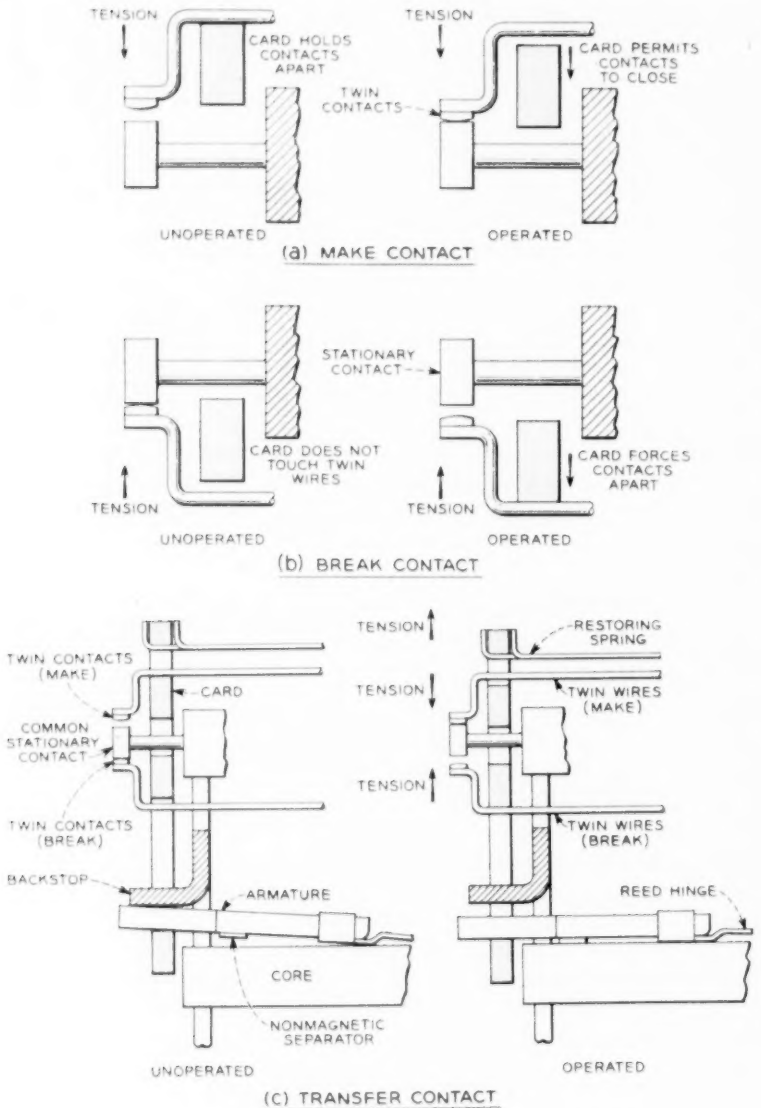


Fig. 4—Principle of contact operation.

4. PRINCIPLE OF CONTACT OPERATION OF THE AF RELAY

The AF relay uses what has been called the "single card system" for actuating the contacts. This is in contrast to other code card systems which require two, three or four coded cards in each relay. The method for obtaining individual make and break contacts with this system is shown in Figs. 4a and 4b, and a means for obtaining transfer contacts, in which both make and break twin contacts are associated with a common stationary contact, is shown in Fig. 4c. As indicated on the figures, the following principles are incorporated in this method of actuation:

1. In general, three basic wire spring assemblies are required. Two of these carry movable twin wires for make and break contacts and are identical except for some details in forming at the terminal ends for convenience in wiring. The twin wire assemblies are mounted on either

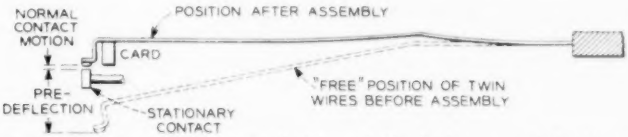


Fig. 5—Contact forces are controlled by relatively large predeflections of the twin wires.

side of the stationary wire assembly, which consists of a group of relatively heavy wires molded into plastic sections, one a short distance behind the contacts and one near the rear of the relay. These sections are rigidly supported in the relay structure.

2. Moving twin contacts on separate twin wires are used with every stationary contact. This arrangement assures good reliability and greater freedom from open contacts in the presence of dust and dirt. In addition, contact chatter is reduced as both contacts must be open simultaneously in order to interrupt the circuit.

3. As shown in Fig. 5, each group of twin wires is tensioned toward the stationary wires by means of large predeflections before assembly, so that the contact forces are determined by this predeflection. Good control of the contact force is assured without need for hand adjustment because small variations in deflection of the low stiffness springs do not result in appreciable changes in force. For this reason, the force is stable and is not appreciably affected by wear of the contacts.

4. The twin wires are actuated by a single punched fiber card. Since the tension in the twin wires is always in a direction to hold the contacts closed, the card serves to hold the make contacts open when the relay is unoperated and the break contacts open when the relay is energized.

5. The card is supported by the armature on one side and a restoring spring on the other. The restoring spring supplies the force to hold the armature against the backstop and to hold make contacts open when the relay is unoperated, while the armature supplies the force to hold the break contacts open when the relay is operated. However, since the armature must also overcome the tension in the restoring spring, the entire spring load must of course be overcome by the pull of the armature.

6. The twin contacts are held in good registration with their associated stationary contacts by means of molded guide slots in the stationary plastic member just behind the card. These guide slots are slightly wider than the diameter of the twin wires so that these wires are free to move in the direction of the armature movement, but are restrained against lateral motion.

7. The close proximity of the card to the contacts is important in minimizing contact chatter and in substantially eliminating locked contacts, i.e., contacts which fail to open because of interlocking of roughened surfaces. The close spacing results in a rigid coupling between the card and contacts, so that the static and dynamic forces associated with the armature and card are available to break loose any incipient lock which might develop.

As the armature moves toward the core, the particular point in its travel at which make contacts close and break contacts open depends upon the dimensions of the card between the surface which bears against the armature and the surfaces which engage the twin wires. By proper selection of these dimensions, any contact can be controlled to operate early or late in the travel as desired. By this means, several sequential contact arrangements may be obtained. For example, if the break contact in Fig. 4c is controlled by the card dimensions to open earlier in the travel than its associated make contact closes, the resulting arrangement is called an "early break-make" transfer. Similarly, an "early make-break" transfer, often called a "continuity" may be obtained by selection of card dimensions which will assure that the make contact closes before the break contact opens. If both contacts operate simultaneously, the result is a "non-sequence" transfer.

From the above it is evident that the card surfaces which engage the twin wires must be in different positions for early contacts as compared with late contacts. This is illustrated in Fig. 6 which shows an early break-make, an early make-break and a non-sequence transfer side by side. Of the contact pairs shown, only two operate early, and this is accomplished by means of steps in the actuating surfaces of the card.

Thus, if no sequences were required, the card would have a single straight surface for makes and another for breaks, and only one card variety would be needed for all combinations of makes, breaks, and non-sequence transfers. Where sequences are needed, however, additional card varieties are required with steps in the actuating surfaces for the early contacts.

In order to obtain a wider variety of the contact combinations including various numbers of make contacts, break contacts, sequence transfers and non-sequence transfers on the same relay, it is necessary to provide a variety of different coded stationary and twin wire assemblies, as well as a variety of cards, some of which are illustrated in Fig. 7. The twin wire assemblies differ as to the number of twin wires provided and in the position of these wires across the width of the molded section.

The stationary wire assemblies are always provided with a full complement of twelve wires in order to support the front molded section, which is held in place by spring tension in these wires. However, only certain of the wires may have contacts at the ends. These stationary contacts consist of base metal blocks with 0.010 inch thick precious metal surfaces on either or both sides as needed for makes, breaks or transfers, and any of the three varieties may be welded to any wire. Thus precious metal is provided only where needed for the particular contact arrangements desired.

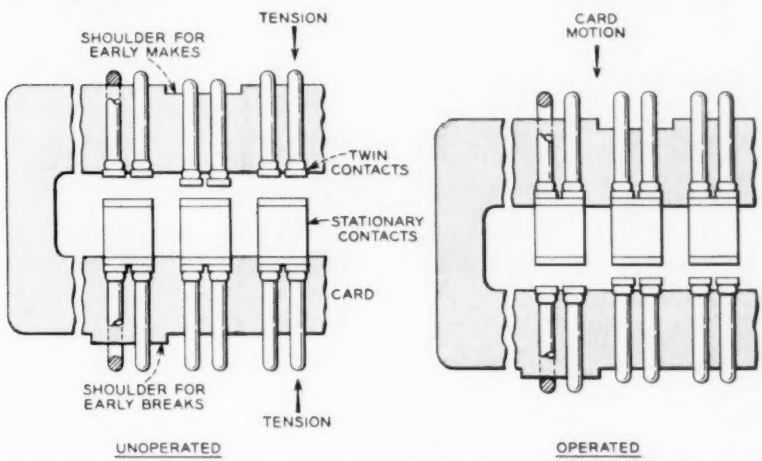


Fig. 6—Early break-make, early make-break and non-sequence transfer contacts, showing how early contacts are obtained by means of shoulders on the actuating card.

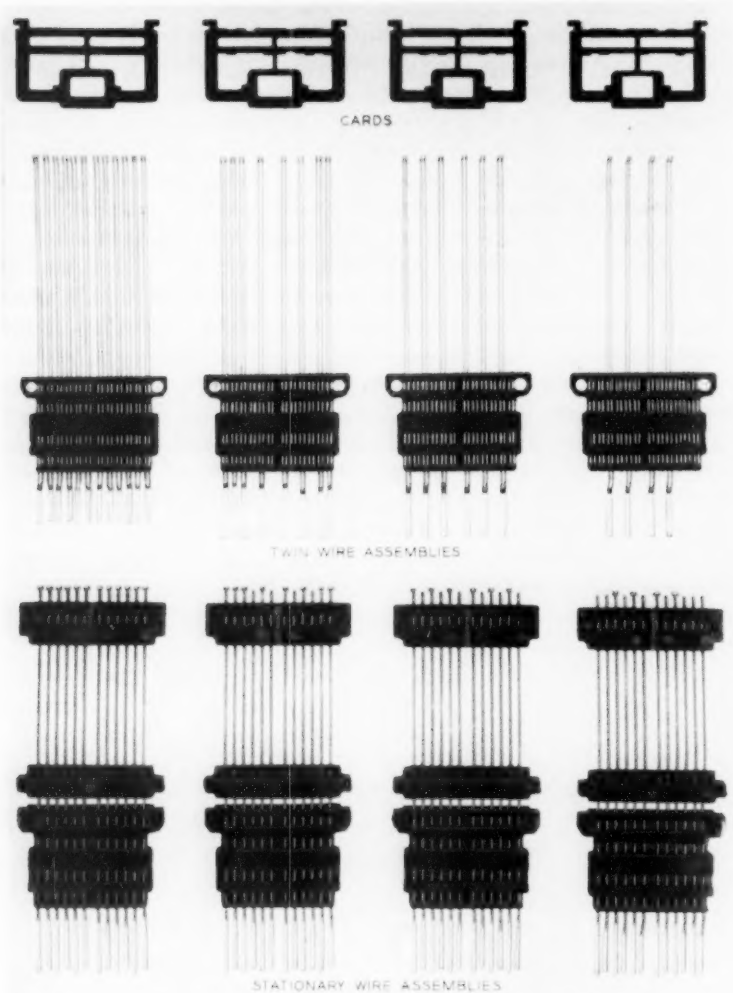


Fig. 7—A few varieties of the coded parts used to obtain various contact combinations.

By using different combinations of stationary and twin wire assemblies with each card variety, a large number of different contact combinations may be obtained. While most of these needed for telephone switching systems use either no sequences at all or a single stage of sequence, a few combinations are provided with "preliminary" contacts. These combinations include two stages of sequence, in which some contacts operate at each of three different points in the armature travel. The preliminary contacts operate earliest in the travel. These are followed by the early contacts of sequence transfers and finally by the late contacts, including ordinary makes and breaks.

To be sure the desired sequences will be maintained during the life of the relays, it is necessary to provide margins in the form of armature travel allowances at each stage. Combinations with sequences will therefore require total armature travels which are longer than those with no sequences, and two stages of sequence will require more travel than a single stage. Accordingly, the AF relay is provided with a choice of three armature travels to correspond with the number of sequences needed. At the card, these travels are 0.026 inch (short) for no sequences, 0.044 inch (intermediate) for one stage and 0.060 inch (long) for two stages.

Thus, combinations including ordinary makes, breaks and non-sequence transfers use short travel. Where sequence transfers are also needed, intermediate travel is used and the early contacts of the sequence transfers operate first. Long travel is used only where preliminary contacts followed by sequence transfers are needed.

5. ARMATURE SYSTEM AND MAGNETIC CIRCUIT

The armature system and the associated magnetic circuits constitute the basic motor element of an electromagnetic relay. The size of the motor element is determined, in part, by the work it must do and here a basic factor is the contact force. On the basis of analytical as well as experimental studies, it was decided to use a contact force of about six grams per single contact, i.e., about twelve grams for the combined force of the twin contacts. Other important factors which react on the design of the magnet are the speed required, winding space, heating,[6] sensitivity, etc. The detailed analysis of the magnetic system and the associated measurements will be covered in separate papers.

The magnetic structure chosen is shown in Fig. 8. The armature is a flat member of U shape which provides desirably large pole face areas.

[6] R. L. Peek, Jr., "Internal Temperatures of Relay Windings", *Bell System Tech. J.*, Jan., 1951, p. 141.

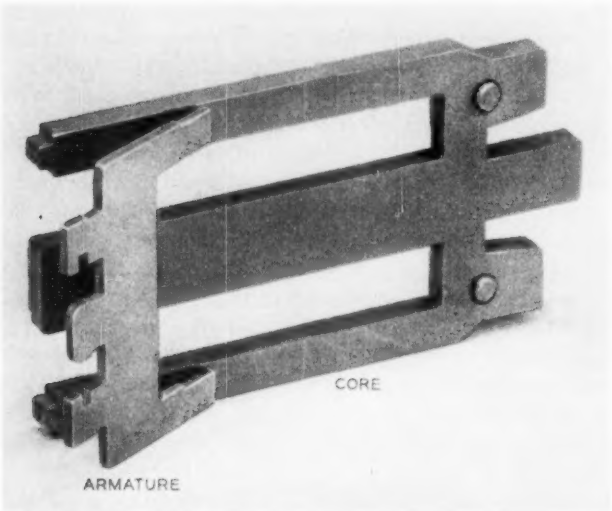


Fig. 8—Magnetic structure of the AF relay.

The core is a simple one-piece E-shape section of sufficient thickness of 1 per cent silicon iron to produce the magnetic flux needed to meet the force and speed requirements and to provide the main member to which all other parts are assembled. The silicon iron has appreciably higher electrical resistance than ordinary magnetic iron and this, together with the rectangular cross-sections of the legs, reduces eddy currents as needed for high speed operation and release. The one-piece construction avoids welded or butt joints common to many magnets. These joints are responsible for added reluctance and hence decrease the magnet sensitivity and require added electrical power to operate a given load. The relatively wide spacing of the legs increases leakage reluctance and, in turn, increases the useful magnetic flux.

After a cellulose acetate filled coil[7] has been assembled to the middle leg of the core, a core plate, shown in Fig. 9, is forced over the ends of the E-shaped core to hold the three legs in good alignment for proper mating with the armature. The core plate also provides the backstop for the armature and serves as a means of gang adjustment of the contacts covered more completely under the Relay Adjustment section of this paper.

The armature is spring supported in a very definite manner to produce

[7] C. Schneider, "Cellulose Acetate Filled Coils", *Bell Lab. Record*, Nov., 1951, p. 514.

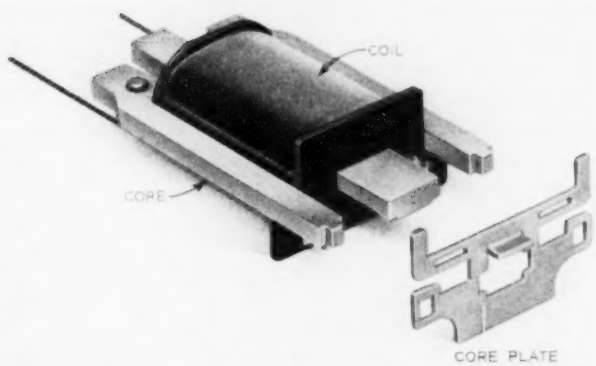


Fig. 9—Legs of the core are held in alignment by the core plate, which is forced over the ends after the coil is assembled.

a minimum of rebound when it is released from its operated position. The conditions for reducing armature rebound were described previously[8] and make it necessary to proportion the forces at the front and rear of the armature properly. The magnitude and the ratio of these forces are a function of the mass distribution of the armature.

The magnet design must not only meet such functional requirements as speed, sensitivity, etc., but it must meet these for several values of armature travel as needed by the variety of contact combinations provided. Another requirement is that the relay be designed to fit on a 2-inch mounting plate and this, in turn, restricts the width of the E-shaped magnet core to slightly less than two inches. The relay is normally mounted with the 2-inch dimension in the vertical direction to allow the contact surfaces to be in vertical planes. The corresponding horizontal dimension in which the relay can be mounted is 1½ inches except for a few special cases. As described in more detail under the section on Relay Performance, the improved magnet design has resulted in a reduction of the magnetic interference between mounted relays to values which are negligible for most practical purposes.

For comparison with the U type relay, the following typical constants of the magnet are of interest (see Table I). The closed gap reluctance, $\mathcal{R}_0$, is the reluctance of the magnetic circuit, excluding leakage paths, with the armature operated and with the iron near maximum permeability. The coil constant, G, is the ratio of the square of the number of turns to the resistance for a full sized coil. The sensitivity, S, is a measure

[8] E. E. Sumner, "Relay Armature Rebound Analysis", *Bell System Tech. J*, Jan., 1952, p. 172.

of the ultimate work capacity of the magnet as related to the power input and has been defined as $S = 5\pi G/\mathcal{R}_0$ ergs per watt.

The favorable low value of closed gap reluctance for the new relay results from adequate cross-sections of magnetic material, the absence of joints, proper mating of the armature and core, and large pole face areas. A low value of reluctance also insures less magnetic interference to other relays and from other relays.

TABLE I

	AF Relay	U Relay
Closed Gap Reluctance $\mathcal{R}_0$, cm^{-1}	0.028	0.065
Coil Constant G, kilomhos	160	160
Sensitivity S, ergs per watt	90,000	39,000

Although the coil constants are the same for the two relays, as can be seen from Table I, the sensitivity of the new relay is more than double that of the U relay, because of the lower closed gap reluctance.

6. MOLDED WIRE SPRING SUBASSEMBLIES

One of the major features of the new relay is the use of molded wire spring subassemblies. Fig. 10 shows a wire spring relay with twelve make contacts, and Fig. 11 shows a comparison of the wire spring assemblies used in this relay and the corresponding parts of the U relay. From this it is clear that the number of parts handled in the assembly of the contact spring members is greatly reduced in the new relay. Not all relays will have twelve contacts and in those cases where fewer contact springs are needed the comparison will not be so unfavorable to the U relay. For six contacts, about one-half of the parts shown will be needed for the U relay, whereas the new relay will again require two wire spring combs. In the new relay three wire spring combs are needed for any contact combination which includes both make and break contacts up to twelve makes and twelve breaks. Four wire spring combs are needed for a relay having twenty-four make contacts.

Two problems had to be solved in providing molded wire spring combs, namely, wire straightening and molding of a multiplicity of wires. Both of these were studied cooperatively at Bell Telephone Laboratories and the Western Electric Company.

Wire is straightened by rotating cam and die members around the unstraightened wire which causes alternating flexing of the wire. For best results, it was found important to shape the cams properly and to

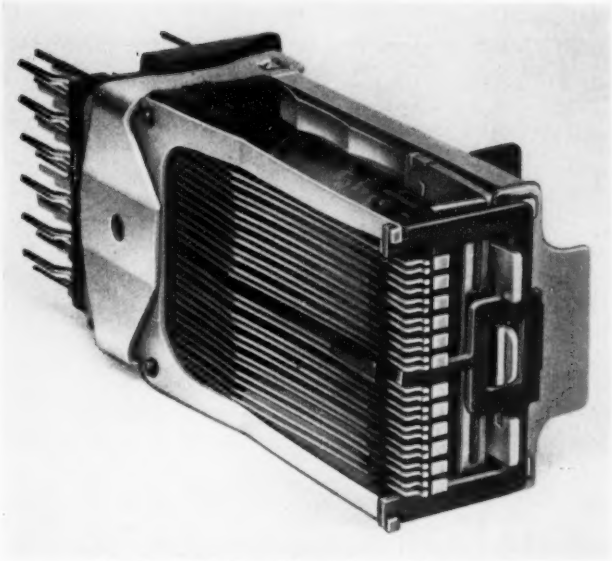


Fig. 10—AF relay with twelve make contacts.

push, rather than to pull, the wire through the rotating cams. By this means it is expected to get straight wire without producing an appreciable twist in it. The Western Electric Company has developed a multiple head wire straightening machine which can be directly associated with the molding press.

A multiplicity of straightened wires is fed into a molding press where plastic molding is used to hold them in proper location. Molding of wire required that the plastic, fed into the die, avoid any appreciable distortion of the wires between unsupported sections. A considerable amount of development work, chiefly by the Western Electric Company engineers, was required to achieve this result. Transfer molding of a thermosetting phenolic plastic has been chosen as the most suitable for producing stable wire spring subassemblies. This is based on the need for stability of the wire positions and because of the ability of the material to withstand the effects of heat. Fig. 12 shows continuous ladders of molded wire spring sections before cutting to length.

The molded sections have a number of features of design importance beyond holding the wires in place. These added features are provided by shaping molded sections to make the remainder of the relay simpler. In particular, these features provide registration pins and holes, guides for

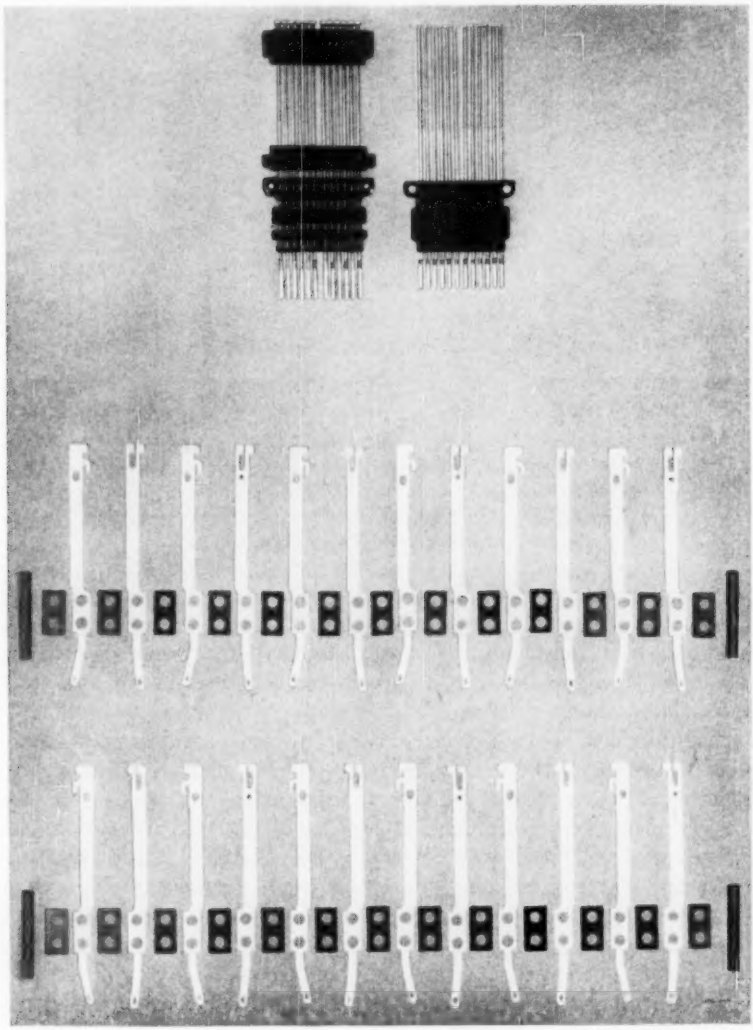


Fig. 11—A comparison of the molded wire assemblies for twelve make contacts with the corresponding U relay parts.

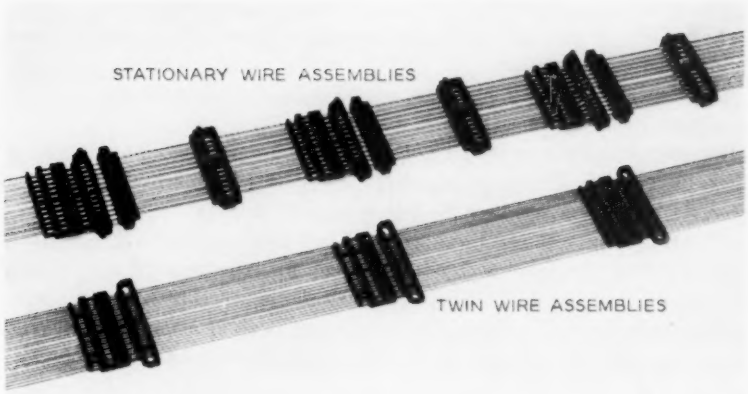


Fig. 12—Molded stationary and twin wire assemblies, before cutting to length.

the ends of the twin wires, cover anchorage, damping material support, etc.

7. CONTACTS AND CONTACT WELDING

Since the primary purpose of the relay is to open and close electrical circuits through the contacts, there has been a special effort to make these contacts as reliable as possible. Accordingly, palladium is used for all contact surfaces. This use of precious metal substantially eliminates opens due to corrosion. Palladium not only gives outstanding reliability but studies indicate that its use results in the best economic balance between manufacturing cost and service because of the reduced maintenance expense.

Open circuits due to particles of dirt between the contact surfaces are largely eliminated by the use of a contact cover, complete independence of the twin contacts described in the section Relay Performance, and the dynamic characteristics of the wire springs. However, to further reduce the incidence of dirt troubles, the surfaces of the twin contacts are coined to a cylindrical shape. This greatly reduces the effective bearing area between the twin contacts and the flat surfaces of the single contacts. Thus, even if an occasional dirt particle should come to rest on one of the contact surfaces, there is small likelihood that it would be in the contact area.

Since welding contacts to wires instead of flat springs is relatively new, considerable attention was given to the development of suitable

techniques. The basic requirements for satisfactory welds are:

1. Sufficient strength to withstand the forces encountered during manufacture and service;
2. Accurate positioning of the contacts on the wires, and
3. Low cost.

These requirements apply to both stationary and twin contacts. However, because of differences in geometry, entirely different methods have been developed for welding the two types of contacts.

The twin contacts are produced by spot welding precious metal contact tapes to the tips of the twin wires. The diagram of the welding circuit is shown in Fig. 13. The condenser C is charged by a power supply to a predetermined voltage. The condenser is then discharged through the primary of the welding transformer T giving rise to the low voltage

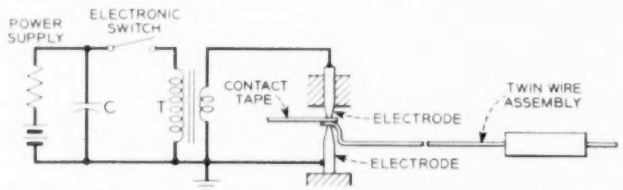


Fig. 13—Diagram showing the essential elements of the spot welding process used for the twin contacts.

high current surge which produces the weld. The contacts are then sheared to length and the surfaces are coined to a cylindrical shape.

The spot welding process did not appear best for welding the stationary contacts to the ends of the wires because of the need to grip the wires with heavy welding electrodes in the limited space directly behind the contacts. Accordingly, a type of welding known as "percussive welding" was developed, which permits one of the electrodes to be placed near the wiring end of the wire springs without developing excessive heat in the wires and which also permits the accurate positioning needed for the contacts in order to control the point of contact closure on the assembled relay. The welding circuit is shown in Fig. 14. The condenser C is charged by means of a direct current power supply, and the condenser voltage also appears on the stationary wire. As the contact to be welded is moved toward the end of the wire, the condenser discharges forming an arc which melts the abutting surfaces of the contact and wire. The constants of the electrical circuit and mechanical system were chosen to assure melting a proper amount of metal at a controlled rate to assure high weld strength. The parts are held together during the very brief

cooling period as the weld is completed. A small resistance R is added in series with the discharge circuit to limit the current and control the arcing period.

That high weld strengths are obtained by this process is indicated in Fig. 13 which shows typical distributions of weld strength for both the percussive-welded contacts and the spot-welded twin contacts. The plots show the percent of contacts with weld strengths equal to or less than any prescribed value within the range of the chart. As shown, the percussive welds are generally stronger than the spot welds which is, in part, due to larger welded areas.

Although percussive welding is more suitable for the stationary contacts welded in the factory, it is planned that occasional replacement of both stationary and twin contacts will be made in the field by spot

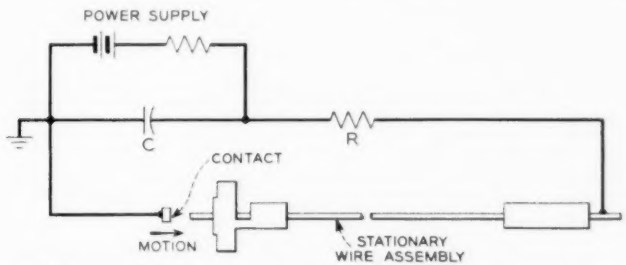


Fig. 14—Diagram showing the essential elements of the percussive welding process used for the stationary contacts.

welding. This will be done with the Bell System field welding equipment[9] provided with suitable electrodes. In this case, however, more expensive all-palladium stationary contacts of special shape would be used to facilitate the spot welding and individual hand adjustment for final position of the contacts will be necessary.

8. STANDARDIZED ASSEMBLY OF CODED PARTS

Since assembly was one of the most promising fields for reducing costs in a new relay design, special effort was made to reduce the assembly cost of the AF relay. Some of the major design features which contribute to low cost assembly are:

1. The continuous molding and fabricating processes for the wire spring subassemblies, which avoid all individual handling of wires and contacts.

[9] W. T. Pritchard, "Relay Contact Welder", *Bell Lab. Record*, April, 1944, p. 374.

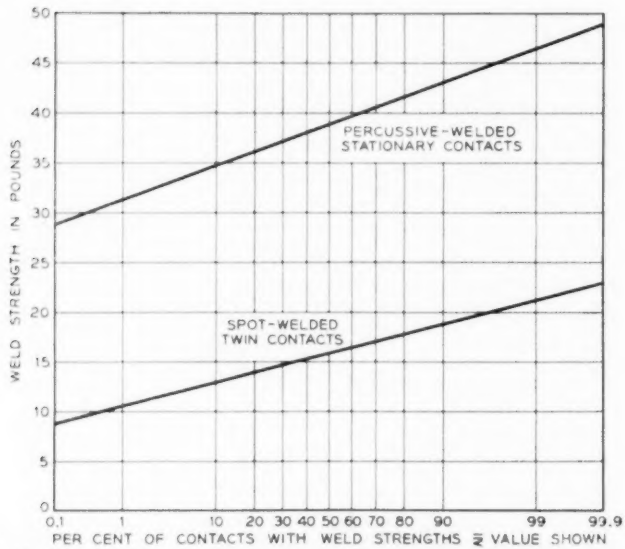


Fig. 15—Typical weld strength distributions for the stationary and twin contacts. The horizontal scale is graduated so that normal distributions will plot as straight lines.

2. Clamping the relay pile-ups by means of a simple spring clamp instead of the more conventional method using screws.

3. A single, easily mounted, operating card.

Less obvious, but equally important is the basic philosophy whereby a large variety of different relay codes are obtained by assembling parts which for assembly purposes are essentially identical for each code. As previously described, the spring combination for each relay is controlled by selection of the proper code card, twin wire assemblies with wires in the proper positions for that combination, and a stationary wire assembly with the right kind of contacts welded to the proper wires. At the present time six different card varieties, fifty twin wire assemblies and seventy-five stationary wire assemblies have been standardized. The twin wire assemblies are provided with any number from one to twelve pairs of wires in various positions while the stationary wire assemblies have from one to twelve contacts in matching positions, with the added variable that each contact may have precious metal on either or both sides as needed. With these it is possible to obtain more than 300 different contact combinations, although only about 100 of these are now needed. Yet, with a few exceptions, each relay code is assembled from

the same number of parts put together in the same manner. By using additional varieties of cards and wire spring assemblies the total number of contact combinations which are possible with the basic design is many times larger than the 300 indicated above.

Other examples of coded parts which are assembled in a standardized manner are the coils, core plates and restoring springs. Although coils vary greatly as to turns, resistance, etc., all are assembled to the cores by the same procedure, using identical spoolheads. The three values of armature travel are controlled by selection of core plates with the proper size of openings, but all core plates are assembled alike. Similarly, the restoring springs are provided in seven varieties including six different thicknesses and seven predeflections to give the desired restoring force, but these variations do not affect the assembly operations.

Standardized assembly of coded parts is of value, not only in reducing the cost of hand assembly operations, but also in providing a more uniform product and as a principle which may make machine assembly practicable.

9. RELAY ADJUSTMENT

Since adjustment expense accounts for a considerable part of the manufacturing costs of older type relays, special efforts were made in the design of the AF relay to reduce the need for adjustment. As a result several types of adjustment used with other relays have been eliminated completely and the remaining adjustments have been simplified. All individual contact adjustment has been eliminated and only two types of factory adjustments are made with the AF relay. These include adjustment of the restoring spring to control armature back tension and a gang adjustment of the stationary contacts to control the points in the armature travel at which the contacts operate. Even these adjustments are needed for only a fraction of the relays as close control of the tolerances in manufacture often causes the back tension and contact operate points to fall within acceptable limits as the relays are assembled.

The gang adjustment of the stationary contacts is made by bending the arms of the core plate, thereby changing the position of the front molded section of the stationary wire assembly which rests on the ends of the arms. Each arm may be bent by means of a screwdriver inserted in the slot as shown in Fig. 16. Rotation of the screwdriver in a counterclockwise direction causes the upper end of the core plate arm to move to the left, carrying with it the upper end of the stationary wire assembly, including the stationary contacts. This reduces the gap between these

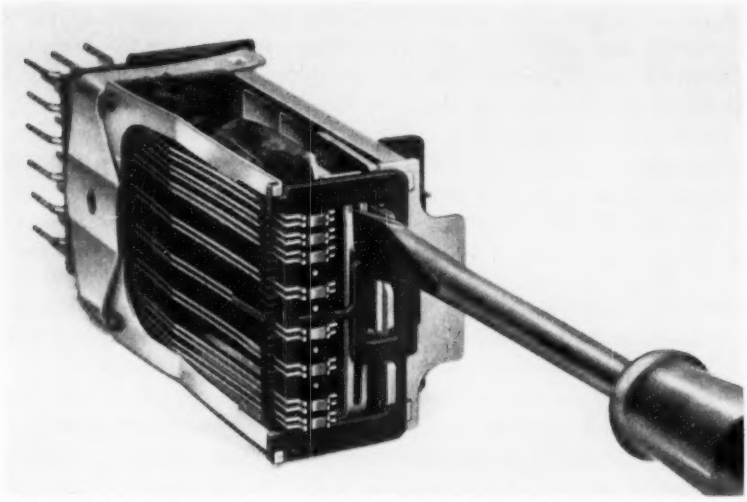


Fig. 16—Contacts may be gang adjusted to operate at the proper points in the armature travel by bending the arms of the core plate with a screwdriver.

stationary contacts and the make twin contacts, thereby causing these contacts to operate earlier in the armature travel. Since the break twin contacts rest against the stationary contacts, these are also moved to the left, reducing the space between the break twin wires and the actuating surface of the card. Thus, bending the core plate arms to the left causes both make and break contacts to operate earlier in the armature travel, while bending the arms to the right causes these contacts to operate later. By bending both arms in the same direction, the operate points of all contacts may be shifted in the same direction. On the other hand, separate arms permit adjustment of the upper relay contacts independently of the lower contacts, thereby increasing the latitude of adjustment.

The parts of the relay are dimensioned so that no adjustment of the core plate arms is required, except to compensate for variations in manufacture of the relay parts. Hence relays assembled from parts made with sufficient accuracy do not generally require adjustment.

The restoring springs may be adjusted for the proper armature back tension by the use of a simple spring bending tool applied to the side arms. However, springs are provided with various predeflections and thicknesses to correspond with various numbers of make twin contacts which must be held open in the unoperated position. Again, no adjustment for back tension is necessary except to compensate for variations

in manufacture, as the restoring spring tension is normally just sufficient to overcome the tension of the make twin wires and hold the armature against the backstop within acceptable force limits. Close control of the tension bends in the wires and restoring springs reduces the frequency with which adjustments are needed and a large portion of the relays do not require this adjustment.

Types of factory adjustment which are common on other relays but which have been eliminated entirely on the AF relay include adjustments for contact force, individual adjustment of contacts for contact operate point, and adjustment for armature travel. Contact force is controlled by means of the large predeflections of the twin wires as mentioned previously. Individual contact adjustment is eliminated by close control of tolerances combined with the single card method of actuation, and by the simpler gang adjustment used when necessary. Adjustment for armature travel is eliminated by the use of close tolerances on the controlling dimensions of the parts.

Adjustments of worn relays in the field may be limited to gang adjustment of the contacts and back tension adjustment as described above. Other adjustments may include burnishing the contacts to remove surface irregularities, replacement of contacts and individual contact adjustment as mentioned previously, and replacement of the card if it should become badly worn or damaged. If card replacement is necessary this may be done without dismounting the relay from the mounting plate and without disconnecting the associated wiring.

10. RELAY PERFORMANCE

As previously stated, the broad objective in the design of the AF relay has been to reduce the annual charges for the use of this relay in the telephone system. Part of this reduction comes from lower manufacturing costs; the remainder comes from savings associated with the improved performance characteristics, such as long life with relatively low maintenance expense, reduced power consumption, and increased speed which reduces the number of units of certain types of equipment, such as markers, needed for telephone central offices. A brief description of some of the principal characteristics of the new relay follows.

Load and Pull Characteristics

Typical load and pull curves for a wire spring relay with twelve early break-make transfer contacts are shown in Fig. 17. The abscissa shows the motion of the armature as it travels from the unoperated position

to the operated position, and back again. This is measured at the centerline of the card and hence is also the card motion. In the unoperated position the armature rests against a backstop, which is part of the core plate. In the operated position it rests against 0.006-inch thick nonmagnetic separators which prevent the armature from touching the core. The ordinate shows the spring load, which opposes the armature motion

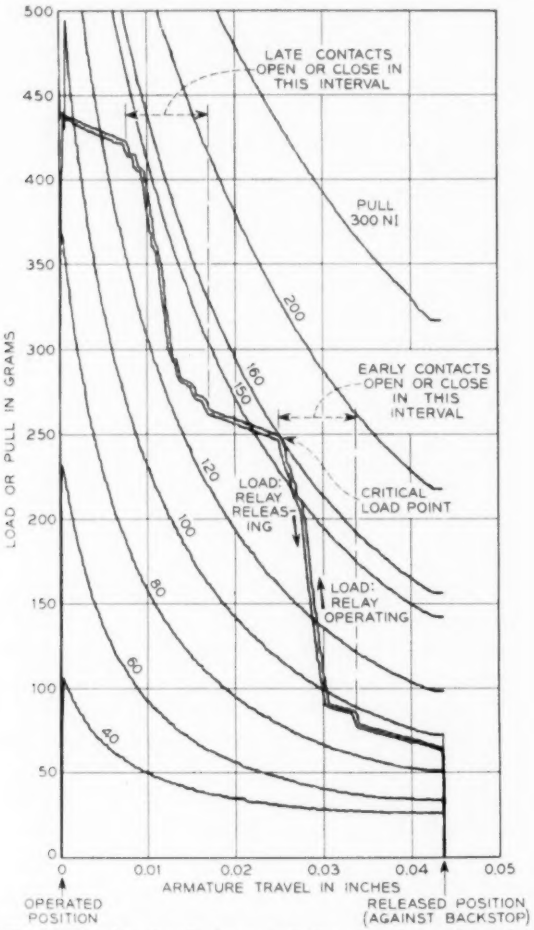


Fig. 17—Typical load and pull characteristics of a wire spring relay with twelve early break-make transfer contacts.

toward the core, and the magnetic pull acting on the armature for various numbers of ampere turns in the winding. These pull and load curves are also measured at the card.

Examination of the load curves shows several features of the relay. The armature back tension, or force, holding the armature against the backstop is about 65 grams in this case. As the armature moves toward the core, the spring load increases along the upper of the two nearly-parallel load curves until it reaches a final value of about 440 grams in the operated position. As the armature is allowed to return to its original position, a second curve, just below the original curve, is obtained. The area between these two curves is a measure of the energy loss due to mechanical hysteresis, or friction, in the relay. As can be seen from the curves, the friction in the new relay is very low and is a small fraction of the spring load at all values of armature travel.

The shape of the load curves is characteristic of AF relays with intermediate travel (0.044 inch). The load increases rapidly in two regions, corresponding to the intervals in which the early and late contacts operate. The rapid increases are caused by the armature and card picking up the additional load of the twin wire springs. Each of the 48 twin wires is picked up almost abruptly at various points and the summation of these additions to the load gives the irregular appearance shown.

The pull curves of Fig. 17 are for essentially static conditions since the armature was restrained to move slowly through its travel while the curves were automatically recorded. These curves are of interest because they show the ampere turns necessary to assure operation of the relay and also values which will assure the armature will not leave the backstop. For example, the "critical load point," or point on the load curve which requires the greatest number of ampere turns, is seen to occur at 0.025-inch travel and 250 grams, which under static conditions would require at least 160 ampere turns in the winding to assure complete operation. On the other hand, as little as 94 ampere turns could cause the armature to leave the backstop and might cause operation of one or two contacts. Hence, a lower value must be maintained to assure that the armature will remain at rest against the backstop. This information is important for relays having non-operate requirements. Similar information may be obtained for limiting ampere turn values which will assure that the armature will remain in the operated position (hold requirements) and, again, which will assure complete release to the backstop position (release requirements).

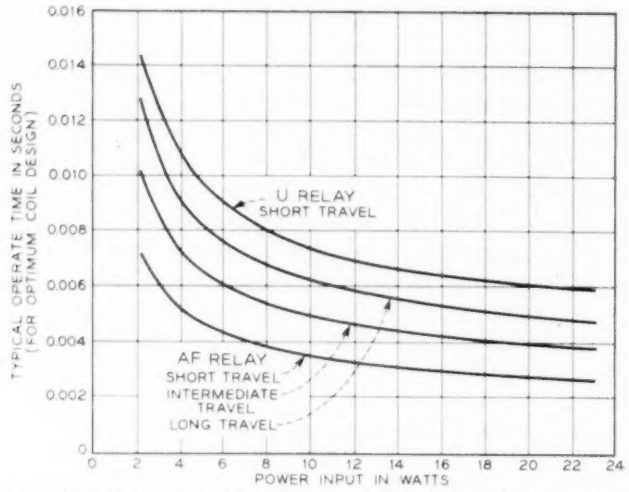


Fig. 18—Typical operate times of speed relays with optimum coil designs for high speed operation.

Speed of Operation and Release

Typical operate and release times of the AF relay are shown in Figs. 18 and 19. Fig. 18 shows the operate times for "speed relays" in which the speed is limited primarily by the time needed to accelerate the mass of the moving system, and is not affected appreciably by the spring loads. These are relays with coil windings of about 1000 ohms resistance, or less, corresponding to power inputs of 2.3 watts, or greater, when connected to a 48-volt supply. For each value of resistance, the operate time shown is obtained with windings having the optimum number of turns for shortest operate time. This time is plotted as a function of power input for various armature travels. Short travel relays have typical operate times varying from about 2.5 to 7 milliseconds as the power is reduced from 23 to 2.3 watts, or as the resistance is increased from 100 to 1000 ohms, using the appropriate number of turns in each case. Intermediate and long travel relays have longer times. For comparison, a short travel U relay is also shown. This relay requires about twice the time to operate as the corresponding AF relay. The improvement with the AF relay is due primarily to the lighter mass of the moving system and slightly shorter travel due to better control of tolerances. Better control is inherent with the single card system for contact actuation and is accomplished without individual adjustment of contact or backstop position, both of which are hand adjusted on U relays.

Typical release times for the AF and U relays are shown in Fig. 19, with time plotted as a function of the number of contact pairs for relays equipped with standard 0.006-inch thick nonmagnetic separators. In this case the improvement is greater than two to one, due principally to the lighter moving parts of the AF relay and lower eddy current effects of the rectangular silicon iron core.

Power Requirements

The nominal power required to assure operation, with some margin, of relays with windings designed for minimum power consumption is shown in Fig. 20. Included is an allowance for adverse variations in magnetic structure, winding and loads. Since least power will be used by the largest coil wound with the finest wire consistent with meeting the ampere-turn requirements for the various contact loads, the curves are discontinuous and have steps as the wire sizes are shifted from one size to the next to meet the ampere turns needed for increasing loads. Again, the corresponding U relay power is shown for comparison. For corresponding numbers of contact pairs, the AF relay requires about half the power of the U relay, except with fewer contact pairs where the power in each case depends upon the use of No. 41 gauge wire. This comparison applies only when coils of optimum design for minimum power are used on both relays. In practice, the coils are selected for the best economic balance between power consumption, cost of the coils and value of speed of operation. Coils designed for minimum power are rela-

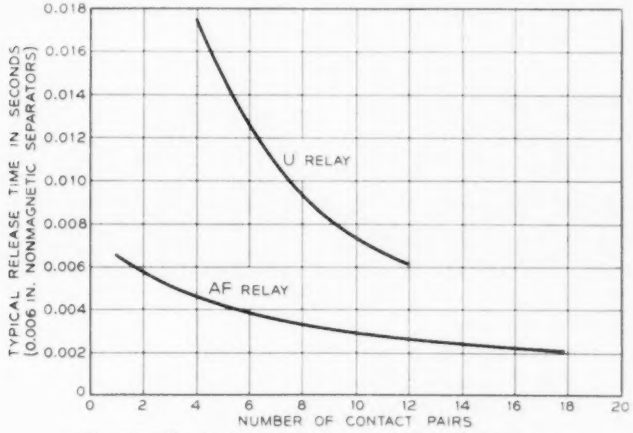


Fig. 19—Typical release times of AF and U relays.

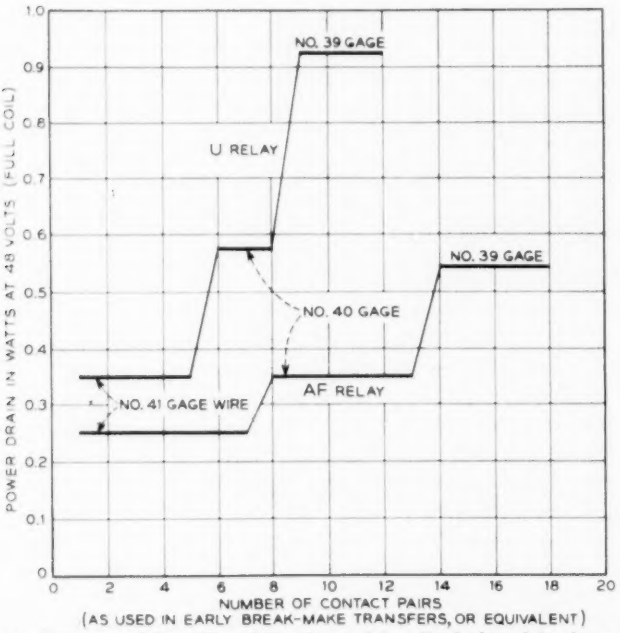


Fig. 20—Power used for AF and U relays with coils designed for least power.

tively expensive because they contain as many turns of fine wire as the available space permits, and their use is economical only on relays which are operated an appreciable portion of the time and where speed is relatively unimportant. The reduced power required for the AF relay is due principally to an improved magnetic structure, shorter travels for similar contact combinations, and lower contact forces.

Contact Performance

The principal characteristics which must be considered in evaluating contact performance include chatter, erosion or wear, susceptibility to open contacts and locking, and changes in these characteristics with wear of the relays. In general, all these features are improved on the AF relay compared with the U type.

Typical chatter on closure of make and break contacts on U and AF relays built for moderate and fast operation is shown in Fig. 21. The degree of improvement of the AF relay is striking. The reduction in chatter has direct circuit advantages in reducing the possibility of false

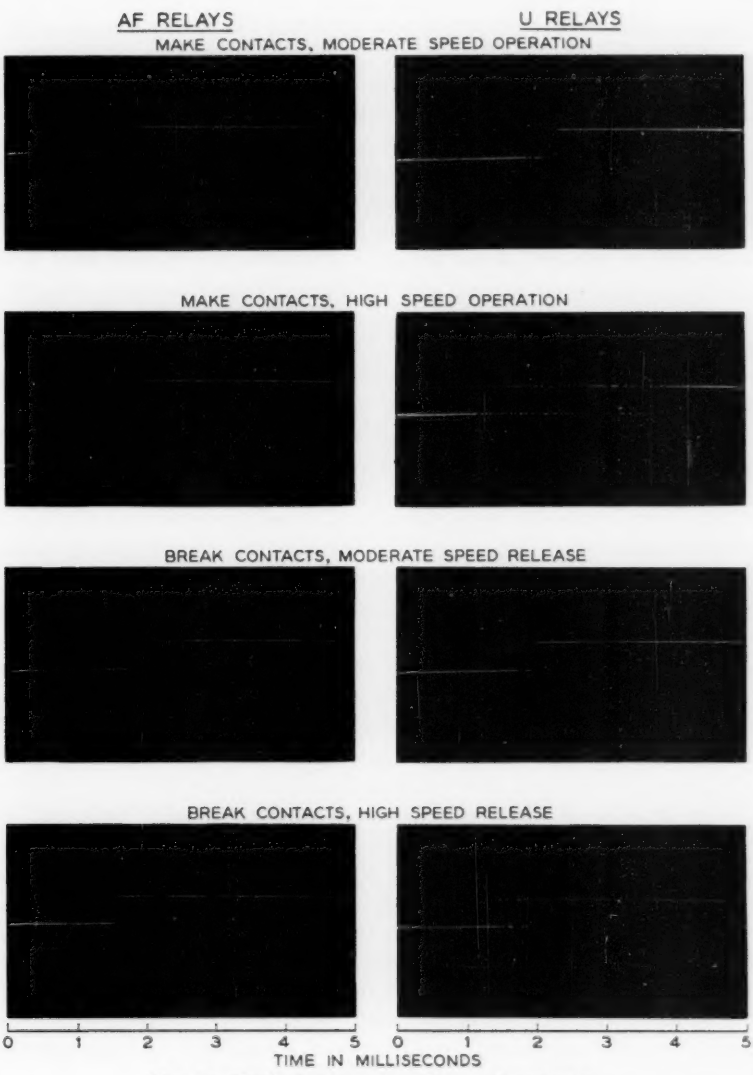


Fig. 21—Typical chatter on closure of contacts.

operation of associated high speed equipment and also indirect advantages in prolonging the life of the contacts. The improvement is due largely to the type of card operation of the completely independent, low-mass twin wires and also to the low mass of the moving system which excites less vibration of the relay structure as a whole. The placement of the card close to the contacts allows the full contact force to be developed within a very short time, and the low mass of the twin wires stores little kinetic energy to cause reopening due to wire vibration.

A particularly troublesome type of chatter occasionally experienced is caused by rebound of the armature after striking the backstop. This chatter is objectionable because of its long duration which is of the order of a millisecond and may occur several milliseconds after the initial opening or closure of the contacts. This increases the possibility of false operation of associated circuits. Accordingly, a fundamental study was made of the means for reducing armature rebound, as previously mentioned. As a result, changes were made in the suspension of the armature and in the position of the backstop which substantially eliminated chatter due to armature rebound.

Electrical erosion of the contacts is reduced on the AF relay because of less chatter and because of the lower energy levels controlled by the contacts, where these are used to operate other AF relays. This improved performance not only reduces maintenance but permits the use of less expensive, smaller size contacts.

Contact locking is substantially eliminated on the AF relay because of the card operation, where the static and dynamic forces associated with the card and armature are available to break loose any incipient lock. Open contacts are reduced by (1) protecting the contacts from dust with a small cover, (2) rounding the twin contact surfaces to reduce the effective areas on which particles must lodge to cause opens and to increase the pressure on the areas, (3) the use of palladium contacts, (4) the dynamic characteristics of the wire springs, and (5) the use of twin contacts on completely independent twin wires. The complete separation of the twin wires is an important feature in reducing open contacts. As shown in Fig. 22, the flat punched springs of the U relay carry twin contacts but these are mounted on tips which are separated by a relatively short punched cutout. This limited separation did not achieve the full advantage of twin contacts as a sufficiently large particle of dust under one contact could cause both contacts to be held open. A subsequent design, known as the UB relay,[3] used a longer cutout, resulting in greater independence of the twin contacts with a significant reduction in contact opens. The AF wire spring relay achieves complete

separation by use of separate twin wires and a significant part of the improvement with respect to contact opens is due to this feature.

The improvements in contact chatter and open contacts become even more evident during the life of the relay. As shown in Fig. 23, the contact forces on U relays diminish rapidly with wear of the contacts resulting in increased chatter, more frequent opens and the need for earlier readjustment. Card operation of wire springs with large predeflections, however, assures substantially constant forces, thereby maintaining the initial chatter-free performance and fewer open contacts.

Life

Tests of relays with contacts protected electrically with resistance-condenser networks indicate that the standard AF relays with winding resistances of 700 ohms or greater will have a life in the order of 250–500 million operations before readjustment becomes necessary. With readjustment, of course, these figures can be increased several times.

Where longer life is essential, special features are used to increase the life. With these features a life, before readjustment, of a billion operations is expected for some relays and, with readjustment, all relays equipped with the special features for long life should be capable of a billion operations.

The special features for long life include vibration dampers attached to the twin wires and the stationary wire assembly as shown in Fig. 24, wear-resistant nickel and chromium plate on the armature, core, and core plate, and a long-wearing alloy for the nonmagnetic separators

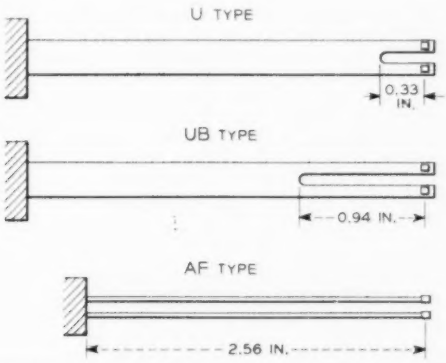


Fig. 22—Independent action of the twin contacts is limited on U and UB relays because both contacts are mounted on a single spring which is notched. The AF relay achieves complete independence by mounting the contacts on separate wires.

welded to the armature. The special features consist largely of variations in finish and material which do not greatly affect the manufacturing processes. The only added parts are the damping members. These are molded from soft but stable polyisobutylene with grooves to receive the twin wires. One damper is attached to each side of the shelf provided on the stationary wire assembly. The twin wires pass through the grooves and are cemented in place. As shown in Fig. 25, these dampers reduce the vibration of the twin wires between the card and the molded section at the rear, thereby reducing the slide between the wires and the card.

Early designs of relays indicated that wear between the twin wires and the card was excessive and that changes in materials would not produce the improvement needed for very long life, particularly with high-speed relays. A fundamental study[10] of the conditions which cause wear was made and it was found that reduction of the sliding motion between the wires and card to 0.001 inch or less was necessary to substantially eliminate such wear. The AF relay meets this requirement. The necessity for such a requirement will be better understood when it is

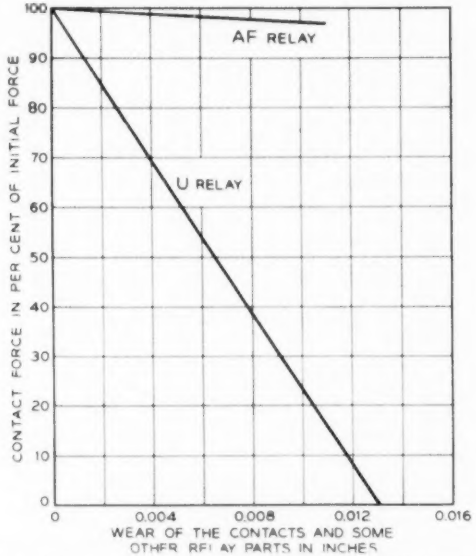


Fig. 23—Contact forces on the AF relay remain almost constant with wear, while U relay contacts lose force rapidly.

[10] W. P. Mason and S. D. White, "New Techniques for Measuring Forces and Wear in Telephone Switching Apparatus", *Bell System Tech. J.*, May, 1952, p. 469.

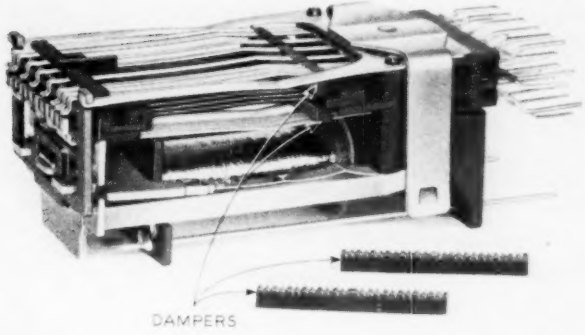


Fig. 24—Where very long life is needed, polyisobutylene dampers are mounted between the twin wires and a molded shelf on the stationary wire assembly.

noted that, for one billion operations, the total slide corresponds to a distance of about thirty-two miles.

Stability

The AF relay is a distinct improvement in stability compared with earlier designs when subjected to shock or temperature and humidity changes. Under severe and repeated variations in temperature and humidity, the largest changes in contact separation are not more than 0.002

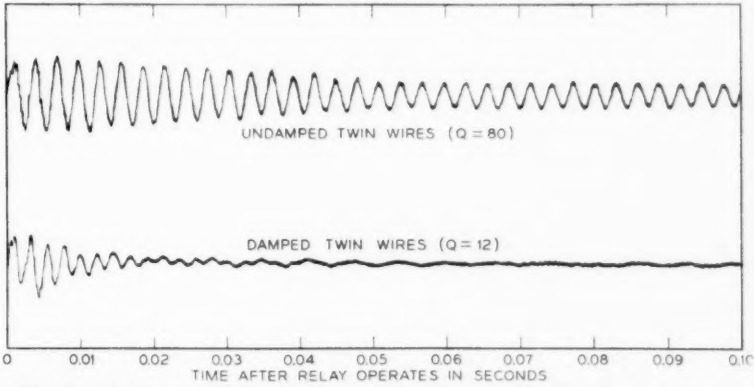


Fig. 25—Oscillograms showing the effectiveness of the polyisobutylene dampers in reducing the vibration of twin wires following operation of the relay. The vibration is measured in the horizontal plane, about midway along the length of the wires.

to 0.003 inch. The improved stability is expected to permit final adjustment and inspection of the relay at the time it is assembled without need for readjustment after it is wired into equipment and installed into service after shipment.

Magnetic Interference

In the past it has often been necessary to maintain large spacing between relays where critical values of current to operate or release the relays must be maintained. In some cases special iron shields were used for further magnetic isolation. Without these precautions, the leakage flux from adjacent relays entered the magnetic circuit of the critical relays and the operate or release currents varied according to whether the adjacent relays were energized.

Magnetic interference between AF relays is substantially eliminated as shown in Fig. 26. This is largely because of the low reluctance of the magnetic circuit resulting from the one-piece core and the large pole face areas between the core and armature. As shown in the figure, the

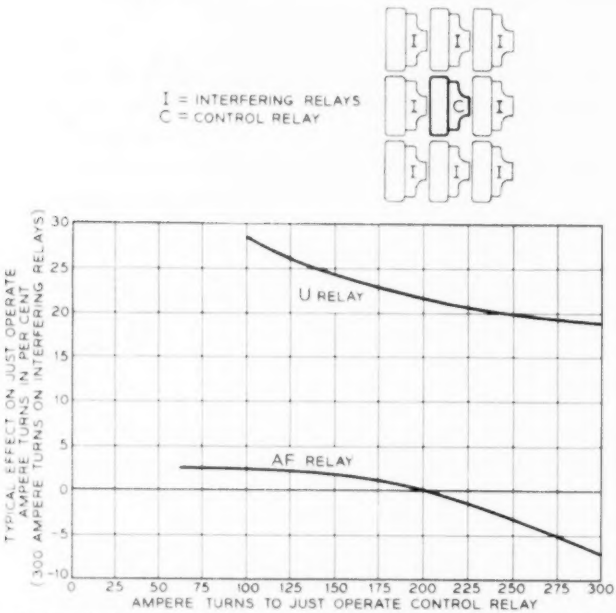


Fig. 26—Typical magnetic interference between AF relays and between U relays, with the relays mounted in the pattern shown.

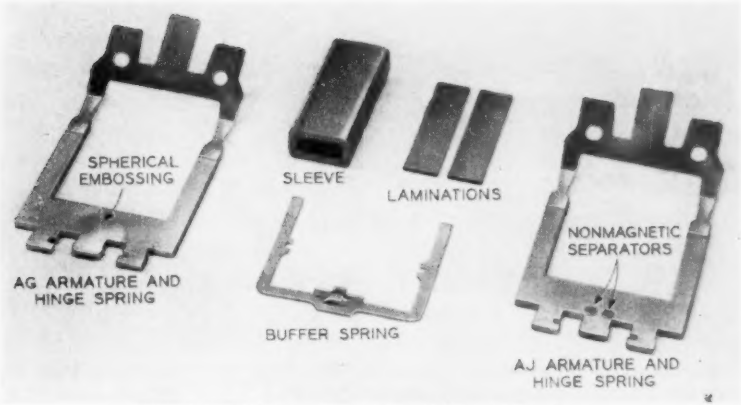


Fig. 27—Additional parts for AG and AJ relays.

measurements were made by surrounding a control relay with eight adjacent closely-spaced interfering relays. The ampere turns to just operate the control relay were varied by changing the mechanical load on the relay, and for each value the change caused by simultaneously energizing the adjacent relays was observed. The improvement of AF relays with respect to the U type is seen to be of the order of ten times for most of the range, with the effects of the adjacent relays being well under 10 per cent up to 300 ampere turns. This is small enough so that no shields or precautions with respect to spacing are required.

11. AG AND AJ TYPE RELAYS

The AG and AJ type relays include modifications of the basic AF design to provide slow release, sensitive, marginal and other additional characteristics. For the most part these modifications are not extensive and the assembled relays closely resemble the AF design.

The additional parts most often used in the AG and AJ relays are shown in Fig. 27. Both relays use thicker armatures with longer side legs than the AF relay, and the armature of the AG relay has a spherical embossing instead of nonmagnetic separators. This reduces the magnetic circuit reluctance of the AG relay when it is in the operated position. In addition, for longer release times, a metal sleeve is assembled over the middle leg of the core, inside the coil. Induced eddy currents in this sleeve oppose rapid changes of flux through the core.

The use of the heavy, embossed armature and a sleeve are sufficient

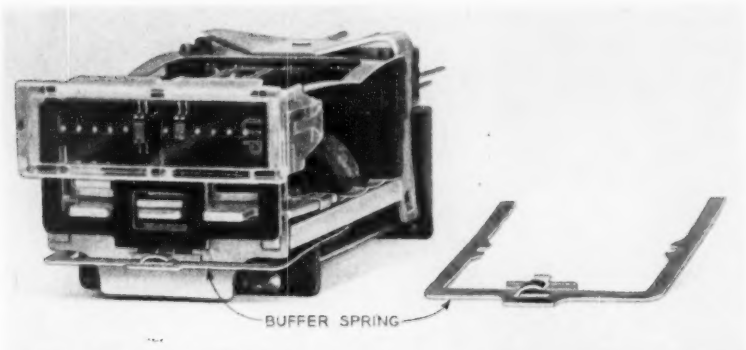


Fig. 28—The buffer spring is used to control the operated spring load, and therefore the release current and release time.

to make the relay slow to release.[11] When the current in the winding of such a relay is interrupted, the flux decays slowly due to the circulating currents in the sleeve. Also, the low magnetic reluctance increases the time for the flux decay by permitting relatively high flux values to be maintained by smaller circulating currents. However, to achieve better control of the release times and to maintain stable adjustment during the life of the relays, the following additional features are used:

1. The cores are annealed in a hydrogen atmosphere, chiefly to stabilize the coercive force of the iron.

2. The core and armature have a wear-resisting chromium plate finish to maintain the nonmagnetic gap between the embossed surface of the armature and the core.

3. The use of a spherical embossing reduces variations in reluctance caused by small angular misalignments between the armature and core.

4. Four sleeves are available including light, medium and heavy copper sleeves and a light aluminum sleeve. These sleeves provide various ranges of release time.

5. A buffer spring is provided on the relay to control the operated load and therefore the release time. As shown in Fig. 28, the buffer spring is normally tensioned against the end of the middle core leg. As the relay operates, however, the card strikes the adjustable tab in front of the middle leg and lifts the spring away from the core so that the spring tension is added to the operated load of the relay. The spring may be adjusted for any desired tension, within limits, and the tab can

[11] H. N. Wagar, "Slow Acting Relays", *Bell Lab. Record*, April, 1948, p. 161.

be adjusted so that the load may be picked up at any desired point in the armature travel.

When a relay is designed for a specified release time, the spread between maximum and minimum times obtained with a particular sleeve is usually greater than desired. Accordingly a sleeve is selected which under normal conditions would produce a somewhat longer time than the specified value. This time is then reduced, as needed, by increasing the buffer spring tension.

Typical release times plotted as a function of the contact load for AG relays with and without sleeves are shown in Fig. 29. Since this figure illustrates release times that are characteristic of the various sleeves, no buffer spring tension is assumed. As would be expected, the heavier sleeves produce the longer release times, which are also greater for relays with fewer contacts. Even the light aluminum sleeve produces several times longer release times than no sleeve at all. For comparison, release times are also shown for the AJ relay, which has a magnetic structure similar to the AG but with nonmagnetic separators in place of the spherical embossing on the armature. The difference between the AJ

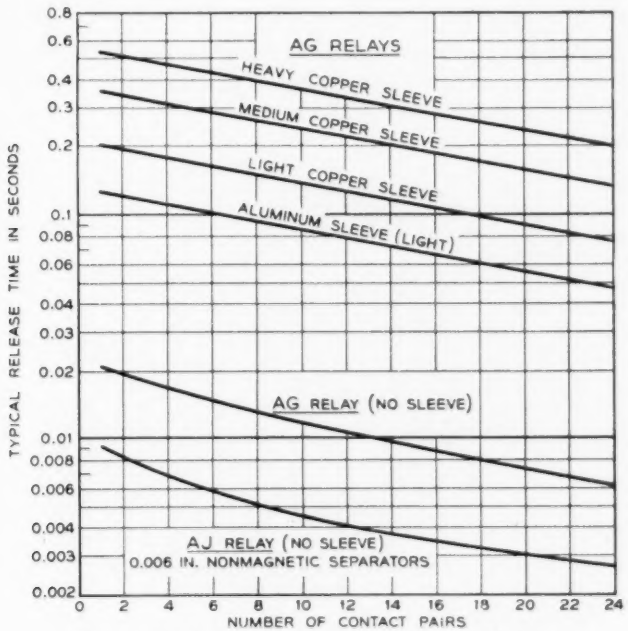


Fig. 29—Typical release times of AG and AJ relays.

release times and those for the AG relay with no sleeve shows the time advantage obtained with the domed armature.

The AJ relay, with its long and relatively heavy armature, is suited for the more critical marginal applications and is capable of operating heavier contact loads than the AF relay. All relays with more than eighteen contact pairs are provided only on the AJ structure. For example, Fig. 30 shows an AJ relay with a full complement of 12 transfers.

A measure of the power requirements for the AJ relay is given in Fig. 31. This shows the power required to assure operation with various numbers of contact pairs for coils designed to consume minimum power at 48 volts. The chart includes allowances for variations in load, magnetic structure and coil, with some margin for changes in these characteristics. Comparison with Fig. 20 shows the power requirements to be slightly lower than for the AF relay except for small numbers of contacts where limitation of wire sizes of the coils is controlling. Under limiting conditions the AJ relay will operate on as little as 0.025 watt.

Other features which may be used to extend the use of the AJ relay for special marginal applications include armatures with various thicknesses of nonmagnetic separators, wire spring assemblies with reduced contact forces, core laminations (strips of iron placed inside the coil, against the middle leg of the core to increase the effective cross-section) and the buffer spring which may be used to control the operated load of the relay and therefore the hold and release currents.

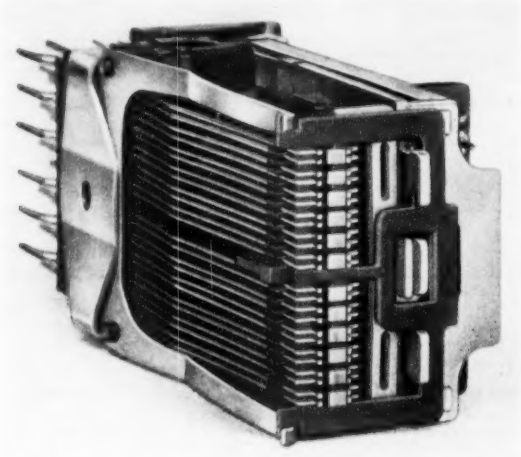


Fig. 30—AJ relay with twelve transfer contacts.

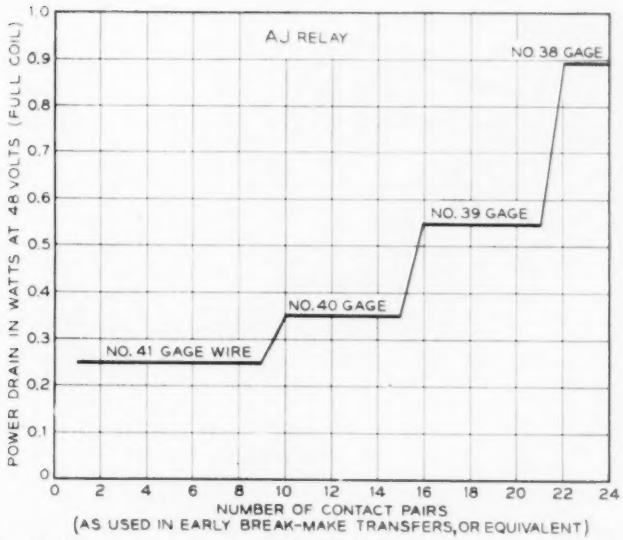


Fig. 31—Power used for AJ relays with coils designed for least power.

A special variety of AJ relay is provided with twenty-four pairs of make contacts as shown in Fig. 32. This relay uses four layers of wire springs and a number of other special parts. As a result, it is often possible to use one twenty-four make contact AJ relay rather than two relays with fewer contacts on each.

12. WIRING THE RELAYS

Connecting wires to the wire spring relay terminals presented a problem which was solved not only for the new relay but by methods which have become important and useful for other apparatus also. The solution came from the invention of a tool, first proposed by H. A. Miloche,[12] for quickly and easily wrapping the connecting wire to the straight end of the wire spring relay terminal.

Early suggestions for making connections to the wire springs included various hooks or bends to simulate the common flat punched terminal having either a hole or notch to facilitate attaching the wires. All of these added some expense to the manufacture of the relay. The added costs were due to two factors (1) forming the wire spring ends required

[12] H. A. Miloche, "Mechanically Wrapped Connections", *Bell Lab. Record*, July, 1951, p. 307.

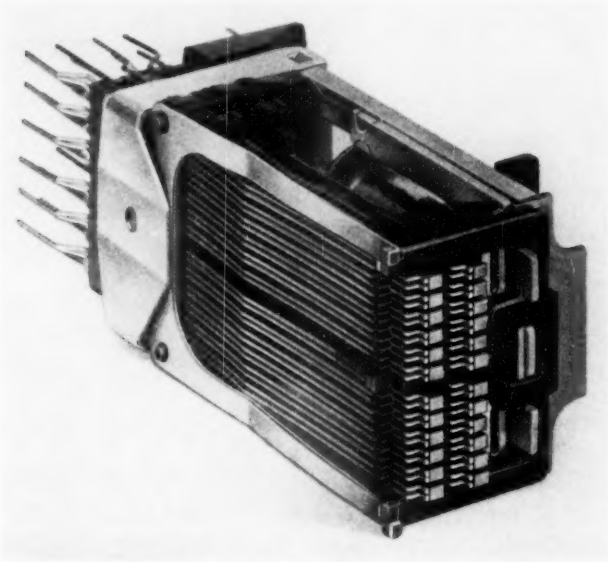


Fig. 32—AJ relay with twenty-four make contacts.

additional operations and (2) the greater flexibility of the wire spring terminals caused some difficulty for the operator so that some increase resulted in the time required to make a connection.

The wrapping tool was first visualized as a simple, trigger-operated hand tool, later as an air or electrically operated tool and still later as a combination tool to do additional operations including cutting and skinning the connecting wire.

Although the wrapping tool was developed to solve a problem which arose in the development of the wire spring relay, it was first applied in commercial practice by the Western Electric Company for wiring to the flat spring relay terminals. Wrapped connections are now used extensively with these terminals, resulting in an improved product at a lower cost. In making wrapped connections to either flat or round terminals, it was the expectation that tinned terminals and wire would be used and soldered together to give a stable, low resistance junction. More recently, however, it has been possible to show that soldering is not required if certain definite dimensional conditions are satisfied by the terminal and the wrapping tool. Solderless wrapped connections will be described in detail in a separate paper.

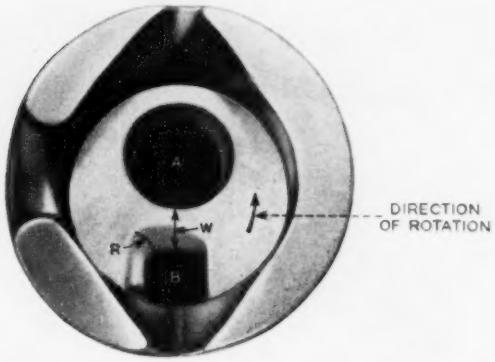


Fig. 33—Muzzle of wiring tool for making wrapped connections.

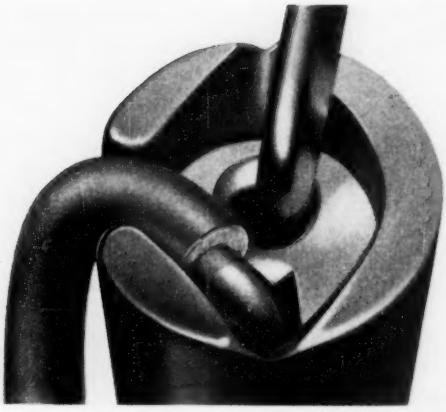


Fig. 34—Muzzle of wiring tool showing terminal and connecting wire in position ready for wrapping.

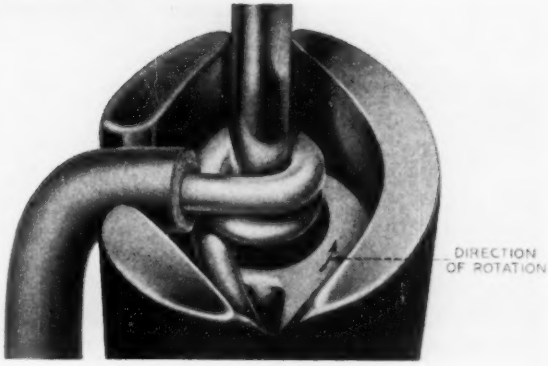


Fig. 35—Muzzle of wiring tool showing two turns wrapped by rotation of the inner cylinder.

Fig. 36—Hand operated wrapping tool for installation and repair service.

Basic Principles of Wrapping Tools

A drawing of the arrangement and action of a wire wrapping tool is shown in Fig. 33. There are a number of dimensions of the tool and of the terminal which have engineering importance. For the purposes of this paper, it will suffice to note that the radius, R, and the wall thickness, W, are important in producing the best wrapped connection.

As shown in Figs. 34 and 35, the inner member of the tool rotates

Fig. 37—Wrapping tool operated by air pressure for factory use.

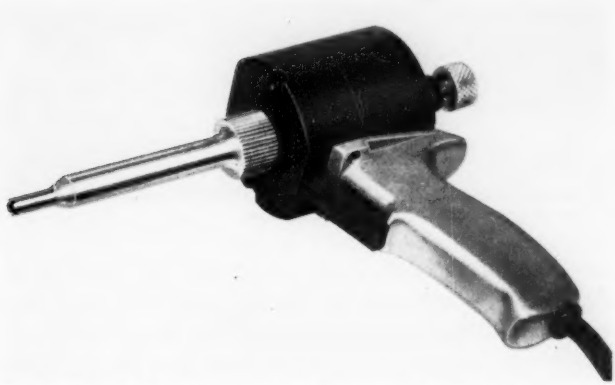


Fig. 38—Wrapping tool operated by electric motor for factory use.

around the terminal which is inserted in hole A of the rotating member. The connecting wire is inserted at B and is anchored by a slight force against the outer stationary member of the tool. As the inner member rotates, the connecting wire is stretched and formed around the terminal until all of the wire length is used. It should be noted particularly that all of the wire is used, making it unnecessary to clip a wire end as in other wiring methods. This is an important detail in avoiding wire clippings which sometimes cause unwanted cross connections in wired equipment units. The tool tip described can be operated by a hand trigger or by motor. Fig. 36 shows a hand powered tool primarily for installation and repair service. Fig. 37 shows a production type tool driven by air pressure developed by the Western Electric Company at the Hawthorne Works. Fig. 38 shows a tool driven by an electric motor used by the Kearny Works of the Western Electric Company. Fig. 39 shows a wrapped connection on a wire spring relay prior to soldering.

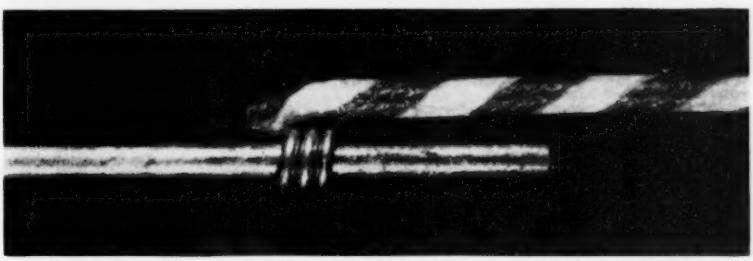


Fig. 39—A wrapped connection on a wire spring terminal, before soldering.

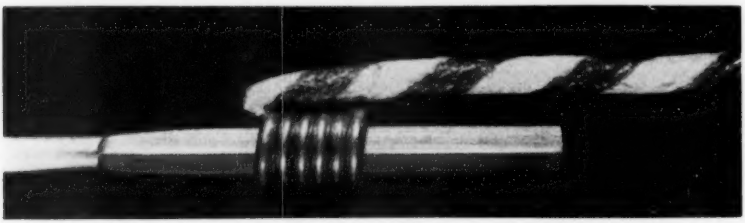


Fig. 40—A solderless wrapped connection on a wire spring terminal.

Another wrapped connection of the solderless type on a similar terminal is shown in Fig. 40. Studies indicate that solderless wrapped connections can be used with a wide variety of materials, including aluminum.

It is interesting to note that a troublesome problem in wiring to the new wire spring relay was solved by the development of a new method which itself has become an important development with broad applications. The wrapped connection with or without subsequent soldering has resulted in better, more uniform and less costly connections made in less time than those made by previous methods.

13. CONCLUSIONS

A description has been given of a new type wire spring general purpose relay for telephone switching systems. Although accurate manufacturing costs will not be available for some time, the new relay is expected to be substantially lower in cost. It provides major improvements in contact performance, reduced power, faster operation and longer life. The new relay also covers a wider field of application than any previous general purpose relay in such characteristics as speed, slow release, marginal operation, number of contacts, etc.

Important economic advantages include lower manufacturing cost of the relay itself and a reduction in switching systems costs resulting from less equipment and reduced power.

The new relay has shown considerable improvement in mechanical life and in contact performance. The life of the relay is of the order of one billion operations. These improvements can be expected to reduce the cost of maintenance of switching systems appreciably.

The development of the new relay has called for a major cooperative effort of many sections of Bell Telephone Laboratories, including such departments as Switching Apparatus Development, Switching Systems Engineering, Switching Systems Development, Research, Materials, Chemical, etc. Without the cooperation of the many members of these

organizations and their special skills, this development would not have resulted in the important and balanced design which has been described. Throughout the development, the associated organizations in the Western Electric Company and the American Telephone and Telegraph Company have made important and guiding contributions, and the New York Telephone Company has cooperated in field trials.

As the spokesman for the project, I wish to express appreciation to the many people who have contributed to this important technical accomplishment. The few names which have been mentioned in the paper were given for historical reasons and cannot replace the large number of important contributors to the project.

To D. C. Koehler, I am indebted for assistance in preparing the paper and the illustrations.

Comparison of Mobile Radio Transmission at 150, 450, 900, and 3700 Mc

By W. RAE YOUNG, JR.

(Manuscript received August 22, 1952)

Based on a series of experiments, a comparison is made of the transmission performance of 150, 450, 900, and 3700 mc in a mobile radiotelephone type of service. This comparison indicates that 450 mc is superior transmission-wise to the presently used 150-mc band in urban and suburban areas. In fact a broad optimum in performance falls in the neighborhood of 500 mc It is concluded that this range of frequencies would be well suited for providing coverage to meet the large scale needs which are anticipated in and around metropolitan areas. Although higher frequencies are less desirable, the tests indicate that 900 mc is somewhat to be favored over 150 mc from a transmission standpoint if full use is made of the possible antenna gain. Above this frequency, transmission performance falls off even assuming the maximum practical antenna gain. Transmission at 3700 mc suffers an additional impairment in that the fluctuations in received carrier level occur at an audible rate as the mobile unit moves at normal speeds. It is concluded that while transmission above roughly 1000 mc for these services is not impossible, it would be decidedly more difficult to employ these frequencies satisfactorily.

INTRODUCTION

From the beginning of mobile radiotelephone services offered by the Telephone Companies, both "general" and "private-line" types, it has been apparent that the number of channel frequencies then allocated for these uses would not be sufficient to meet the service needs in the near future.

The bulk of these needs will be for service in urban and suburban areas, where business activities are concentrated. These areas are now served on a few individual FM channels in the vicinity of 150 mc. However, a larger number of channels, needed to meet anticipated demands and to develop a more efficient system, are not to be found in the

150 mc region. This space is already allocated fully and permanently to a variety of other services. In fact, this situation extends up to about 400 mc. The larger number of channels for these services apparently will have to be found, therefore, above 400 mc.

However, it is essential to know whether these higher frequencies would be suitable for urban mobile telephone service, or whether there exists an upper limit to the suitable frequencies. In order to answer these questions, a series of tests has been made to compare the adequacy of coverage that could be provided at several representative higher frequencies. These tests were conducted in and around New York City. This location is considered to be typical of the larger metropolitan areas.

THE PROBLEM OF EVALUATION

It became apparent early in the tests that it would neither be practical nor accurate to compare service results for the different frequencies by the method of determining the coverage at the various frequencies, and then comparing these. This would have required, among other things, that "coverage" be defined precisely and then measured accurately in order to determine the differences with the desired accuracy.

Instead, it was recognized that commercial coverage is at present considered to extend into areas wherein a small percentage of the locations will have less than commercial grade of transmission. This might be ten per cent, for example. It was further recognized that, while there existed a trend of performance with frequency, comparative tests at any one location showed variations from that trend. Thus, even if transmitter powers were adjusted so as to offset the transmission effects of that trend, performance at any location would not be equal at all frequencies. But while one frequency might give relatively poor transmission in one location, it might give good transmission at another location, etc. Thus, while the locations of poor transmission were found to be different at the various frequencies, the number of such locations would be the same at all frequencies, provided the trend had been offset by adjustment of transmitter power.

Viewing the problem in this way, it was sufficient to test at enough locations in representative territory to establish this trend in a statistical manner.

Other problems in evaluating differences in suitability of different frequencies lay in how to take into account differences in practical antenna gains and differences in frequency stability. These will be discussed in the next sections.

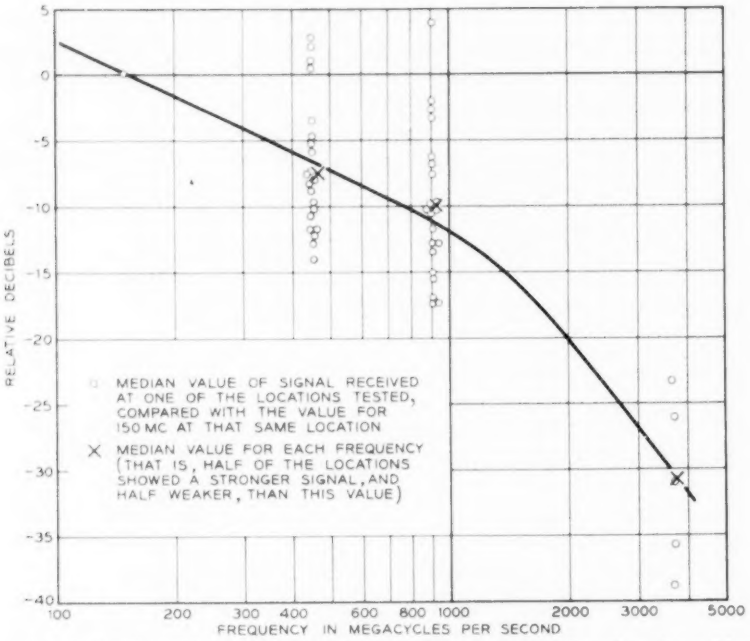


Fig. 1—Median values of received signal power at suburban locations. (Assumes the same power at all frequencies radiated from a dipole and received on a quarter-wave whip.)

OVER-ALL RESULTS

The results of many measurements of path loss between a land radio transmitter and a mobile receiver establish a trend of loss increasing with frequency. This is illustrated in Fig. 1 by the "crosses" which show the strengths of the received signal at higher frequencies as compared with those at 150 mc. The derivation of the values given by the crosses will be discussed in a later section. In the other direction of transmission it appears justified, based upon reciprocal relationships, to assume that path losses from mobile transmitter to land receiver will follow the same trend.

However, although the received signal is seen to decrease with frequency, the amount of received signal which is required to produce satisfactory communication also changes with frequency. The median level of signal required at a mobile or land receiver at various frequencies to override RF noise is given in Fig. 2. The dots here represent the average of many measurements.

Transmitter power required to achieve the same service result at various frequencies has been derived by taking into account the changes of path loss with frequency and also the changes of signal required with frequency. Fig. 3 shows the amounts of power that are required in order to achieve the same coverage in all cases as is now obtained at 150 mc with 250 watts of land transmitter power radiated from a dipole. As shown, the use of an antenna having gain can appreciably lower the land transmitter power that is required. The mobile transmitter power is much less than required of a land transmitter due to the assumption that there are six land receivers located appropriately in the coverage area, rather than just one.

It is apparent from Fig. 3 that the required transmitter power is a minimum in both directions of transmission at around 500 mc. It is also apparent that above this point the required transmitter power increases rapidly with frequency.

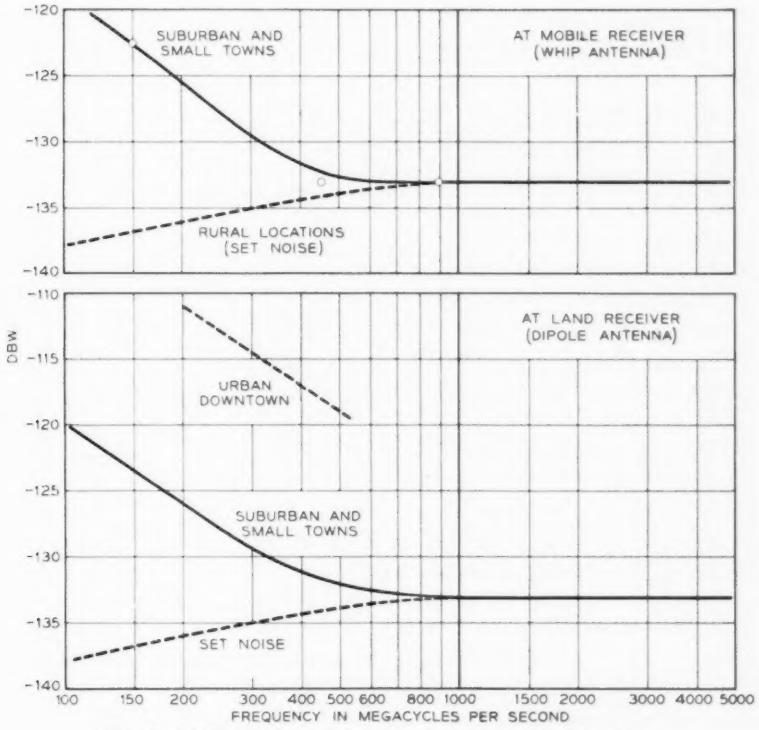


Fig. 2—Median value of signal required to over-ride noise.

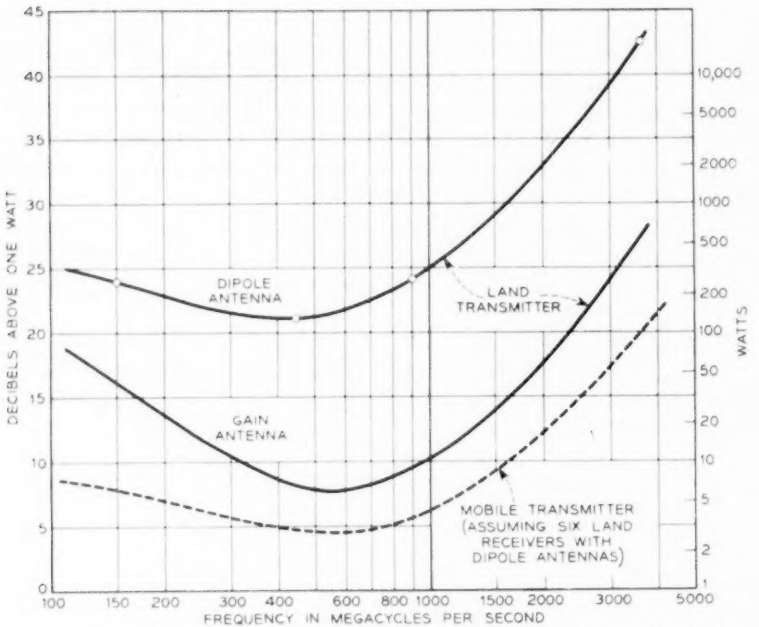


Fig. 3—Transmitter power at antenna input required for urban and suburban coverage. (Mobile antennas are assumed to be quarter-wave ships.)

A word of explanation is needed at this point about the gain antennas which were assumed in one of the curves of Fig. 3. These are antennas which tend to concentrate radiation toward the horizon in all directions. Limits for the amount of gain were based upon the considerations (1) that a set of radiating elements greater than about 50 feet in extent would be impractical to build for this service, and (2) that the vertical width of the beam should not be less than about 2 degrees in order that valleys and hilltops will be covered. The amounts of gain possible within these limits are as follows:

Frequency mc.	Gain-db
150	8
450	13
900	15
3700	15

The mobile antennas were assumed to be quarter-wave whips or the equivalent.

Use of gain antennas for the land receivers would result in still further lowering the required mobile transmitter power. This is not shown on Fig. 3 because the amount of reduction cannot be accurately stated on the basis of present knowledge. It appears certain that the reduction will be at least equal to the antenna gain, and may be appreciably more than this, as indicated later.

The system modulation and pass-band were assumed in the above discussion to be the same at all frequencies. This would not be realistic if the tolerance allowed for frequency instability were a fixed percentage of operating frequency. It may be justified, however, because the necessity for frequency economy and for best transmission performance demands better percentage stability at higher frequencies.

A spot check of transmission, observing circuit merits by listening, has been made to determine the validity of the above results in a very general way. Land transmitter powers were adjusted so that the equivalent dipole power at 450 mc was 3 db less than at 150, and power at 900 mc was 1 db less than at 150 mc. This approximates the powers shown on the "dipole" curve of Fig. 3. The map of Fig. 4 shows the results of this test. While the comparison of circuit merits generally shows a preferred frequency at any given location, the performance appears to be about equal when all locations are considered.

TEST EQUIPMENT ARRANGEMENTS

Tests of transmission outward from the land transmitting station were made on signals radiated from antennas on the roof of the Long Lines Building, 32 Avenue of the Americas, New York City. These antennas were 450 feet above ground. One of the existing Mobile Service transmitters served for the 150-mc tests. Special experimental transmitters were set up for the 450, 900, and 3700-mc tests. All were capable of frequency modulation.

The mobile unit was a station wagon equipped to receive and measur- signals at the various frequencies. The receiving equipment was arranged for rapid conversion from 150 to 450 to 900 mc. The bandwidth (about 50 kc) and system modulation (± 10 kc) were identical at all three frequencies (equal to the existing standards at 150 mc). The 3700-mc tests were handled separately. It was not possible to employ the same bandwidth and deviation, but this does not invalidate the comparison of signal propagation at the various frequencies.

A most useful tool in making these measurements was a device known as a "Level Distribution Recorder", or simply "LDR". This was built

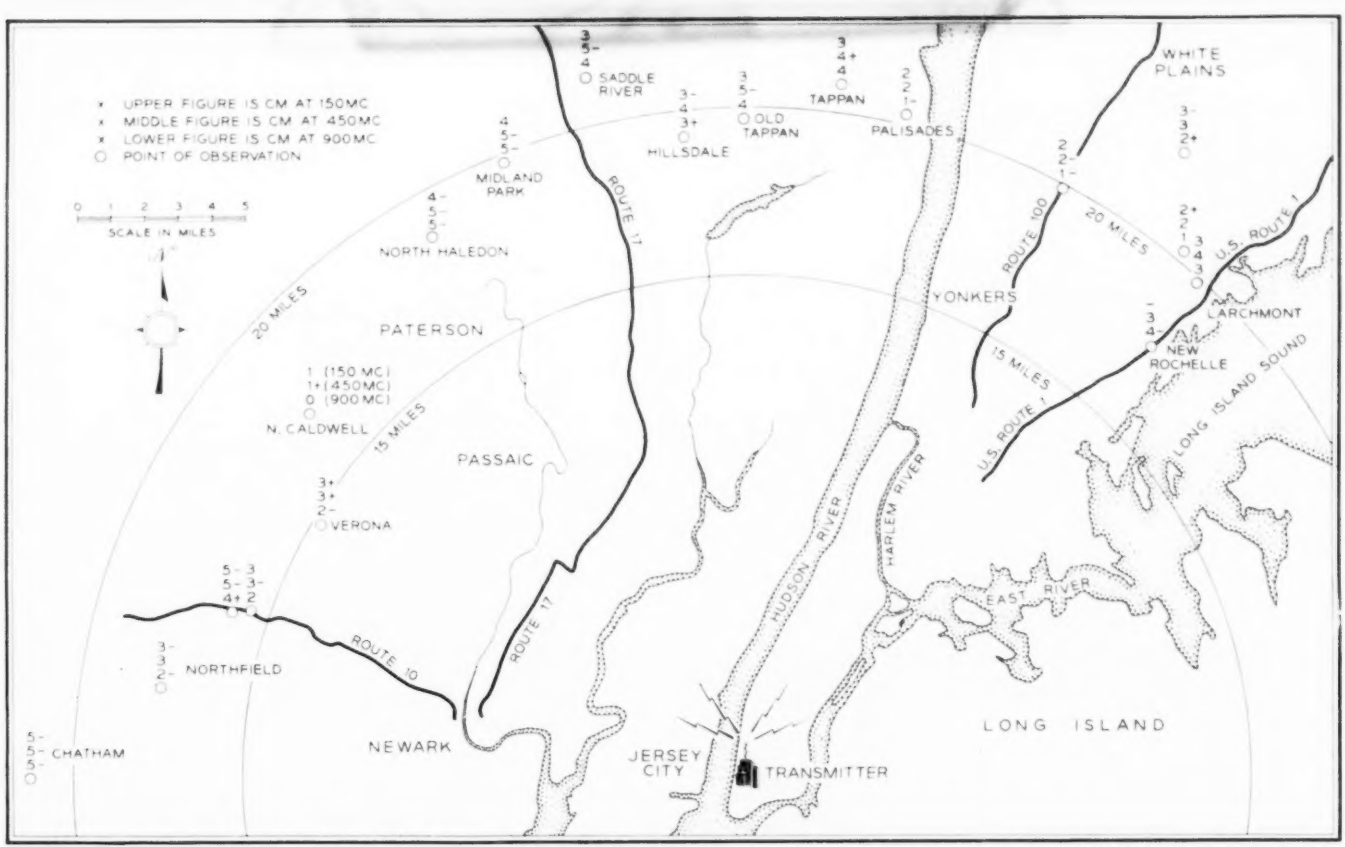


Fig. 4—Circuit merit observations in New York suburbs. Radiated transmitter power was as follows: At 150 mc, 250 watts to a dipole; at 450 mc, the equivalent of 125 watts to a dipole; and at 900 mc, the equivalent of 200 watts to a dipole. Circuit merit grades are as follows: 0—Undetectable; 1—Hopeless, signal unintelligible; 2—Poor, signal understood with much repetition; 3—Fair (commercial); 4—Good, impairment noticeable but not objectionable; and 5—Excellent, no discernible impairment.

especially for these tests and is similar to its forerunners which have been used in the past for measuring atmospheric static noise. The LDR, in combination with a calibrated radio receiver, is capable of taking as many as twenty instantaneous samples of radio signal strength per second, sorting the samples by amplitude, and rendering information on a "batch" of samples from which a statistical distribution curve can be plotted. The LDR was also used for measuring the statistical distribution of audio noise in the output of the radio receiver. The LDR was, in this case, associated with a special converter whose characteristics resemble those of a 2B noise measuring set.

No arrangements were made for measuring radio propagation from mobile unit to a land receiver. It was felt that the comparison by frequencies would be substantially the same as in the outward direction of transmission. It does not follow, however, that the background electrical noise, against which an r-f signal must compete, will be the same at mobile and land receivers. Strength of r-f signal required at land receivers for satisfactory transmission was measured at several typical locations.

RECEIVED R-F SIGNAL STRENGTHS AND PATH LOSSES

The first factor in evaluating mobile radio transmission is the strength of the r-f signal which is received. This is inversely related to the loss in the r-f path. The mobile units of a mobile system are either moving around or, if stationary, are located at random. Since the effects of the many geographical features, buildings, and the like, which influence propagation can combine differently for different locations of a car, even where the locations are only a fraction of a wavelength apart, the only meaningful measure of signal strength is a statistical one. Such statistical answers were obtained by making and recording many instantaneous samples of field strength with the aid of the LDR, mentioned above.

It is of interest to note that whenever the sample measurements were confined to a relatively small area, say 500 to 1000 feet or less in extent, the amplitude distribution of these samples tended strongly to follow along the particular curve known as a Rayleigh distribution. Such a curve and a typical set of experimental points are shown in Fig. 5. The same distribution was obtained at all of the frequencies tested, including 3700 mc. The rapidity of signal fluctuation, as the car moved, was proportional to frequency, but this does not affect amplitude distribution. Such a distribution could have been predicted if it had been postulated that the transmitted signal reached the car antenna by many paths having a random loss and phase relationship. It is thus inferred that in general the signal reaches a car by many simultaneous paths.

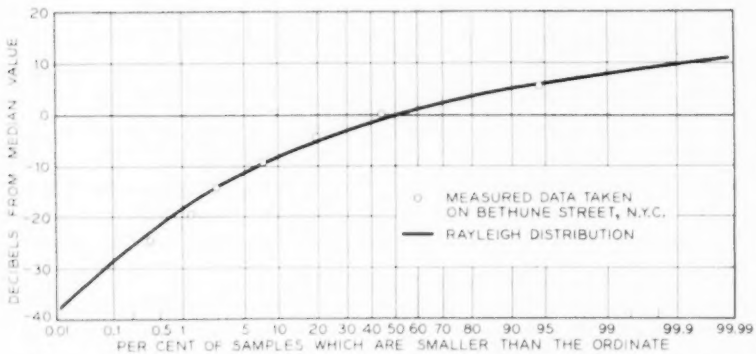


Fig. 5—Typical distribution of test samples of r-f signal strength taken over a small area.

With the shape of the distribution known, only one other value need be given in order to specify the propagation to such a small area. This might be the median, the average, the rms, or any single point on the curve. The one used most often here is the median, that is, the value which is larger than 50 per cent of the samples and smaller than the other 50 per cent. Measurement of the median value by this statistical method was found to be accurately reproducible, and therefore is presumed to be reliable. Successive batches of 200 samples each, all covering the same test area, yielded median values which differed not more than 0.5 db when none of the conditions changed; i.e., transmitter power, antenna gain, and receiver calibration remained the same. This accuracy may seem surprising when it is realized that individual samples differ frequently by 10 db, and often as much as 30 to 40 db.

It was presumed at the outset of the tests that the different frequencies would exhibit different propagation trends with distance. For this reason the samples have been grouped by distance. In presenting these results, it was convenient to express the measurements of received RF signal in terms of path losses. By this it is meant the loss between the input to a dipole antenna at the transmitter and the output of a whip antenna on the test car. These path losses will have, of course, the same distribution as the received r-f signal.

The results of the path loss measurements are given in Figs. 6, 7, and 8 for 150, 450, and 900 mc respectively. These values represent the loss between the input to a half-wave dipole antenna at one end of the path and the output of a quarter-wave whip at the other end. They are shown here as a function of distance from the land station. For distances under

ten miles the data are the result of tests in Manhattan and the Bronx. For each distance a test course was laid out approximately following a circle with that distance as a radius. The data for ten miles and greater distances were obtained on two series of tests along radials from the land transmitter, one of which followed Route 1 through New Rochelle, N. Y., and the other followed Route 10 toward Dover, N. J. For reference, a curve has been given on each of these figures which shows the computed loss based upon the assumption of smooth earth.

A curve labeled "1 per cent" means that in one per cent of the sample measurements the loss was less than that indicated on the ordinate. The meaning of the labels on the other curves is similar. The curve labeled "50 per cent" is, of course, the median.

It will be apparent that the assumption of smooth earth is not applicable to the area tested. The data for median losses are in the order of 30 db greater than the value computed over smooth earth. This additional loss may be thought of as a "shadow" loss arising from the presence of many buildings and structures.

The distribution of losses given in these three curves is wider than the Rayleigh distribution of Fig. 5. This is because the data for each

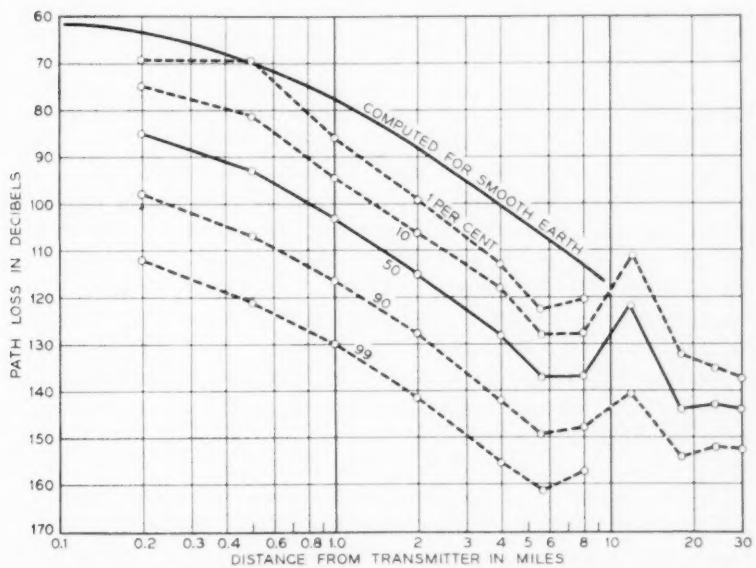


Fig. 6—Measured path loss at 150 mc in Manhattan and the Bronx and suburbs. (Note: Data for 10 miles and greater were taken on Route 1 toward New Rochelle and on Route 10 toward Dover.)

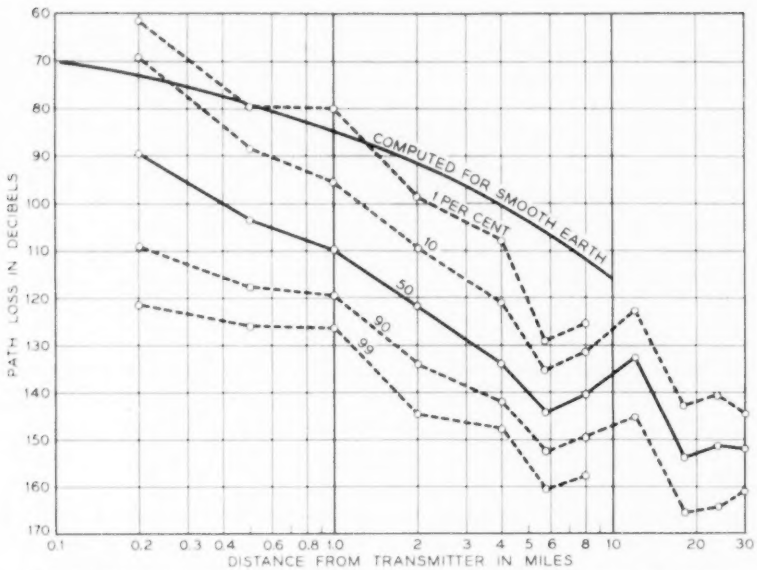


Fig. 7—Measured path loss at 450 mc in Manhattan and the Bronx and suburbs. (Note: Data for 10 miles and greater were taken on Route 1 toward New Rochelle and on Route 10 toward Dover.)

distance are a summation over many different locations rather than a set of samples covering one location.

The data for ten miles and further from the transmitter were taken on routes through suburban areas. The losses at twelve miles appear to be less than the average trend indicated by the curves. This is because data taken at the top of the First Orange Mountain weigh heavily at this distance. It is of interest to note that the losses at distances of ten miles and over are 6 to 10 db less than might have been predicted from the trend at smaller distances, where the measurements were made in city areas. This probably reflects the fact that there is a considerable difference in the character of the surroundings, such as height and number of buildings in the suburban territory as compared with the city itself.

The median curves of loss have been replotted for three frequencies on Fig. 9. This permits a better comparison with frequency. Except very close to the transmitter, the performance at the various frequencies seems to differ by an essentially constant number of db, while exhibiting the same trend with distance. The similarity between frequencies is appar-

ently much greater than the similarity between the median value and the value computed over smooth earth for any given frequency.

It was not possible to get complete enough data to plot a curve for 3700 mc similar to the ones mentioned above. The test setup at this frequency was limited by transmitter power and receiver sensitivity. Only those locations for which path loss was relatively low could be tested. A comparison of results at these locations is given in Figure 10. The curves labeled "1 mi.", "2 mi.", and "4 mi." for Manhattan are the median values obtained along test routes which followed circles of 1, 2 and 4 miles radius from the transmitters. The other curves refer to selected small areas at greater distances on the Hutchison River Parkway and New Jersey Route 10, as indicated. Although the data at 3700 mc not extensive, the trend with frequency seems clear.

More specific data for path losses measured along the routes toward Dover and New Rochelle are given in Fig. 11. Each value plotted here is the median of about 200 samples taken in a small area at the distance indicated. The strong effect of the First and Second Orange mountains at fourteen and sixteen miles on the Dover route is of interest.

The coverage desired in these mobile telephone systems extends into suburban locations. It follows that a comparison of coverage by the

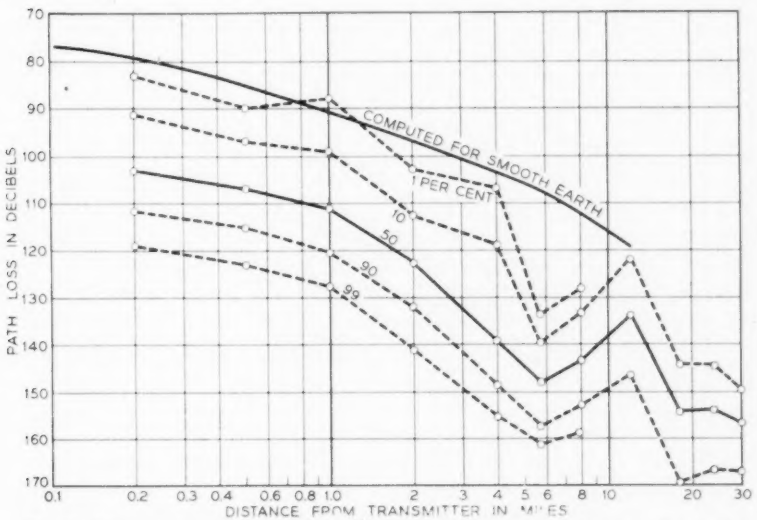


Fig. 8—Measured path loss at 900 mc in Manhattan and the Bronx and suburbs. (Note: Data for 10 miles and greater were taken on Route 1 toward New Rochelle and on Route 10 toward Dover.)

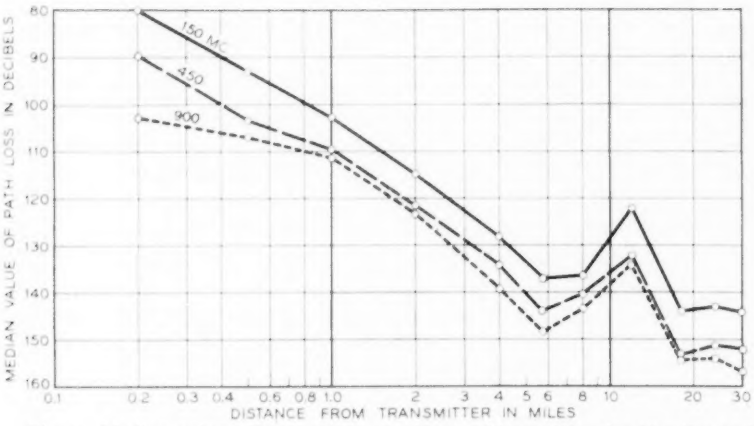


Fig. 9—Median values of measured path losses. (Note: Data for 10 miles and greater were taken on Route 1 toward New Rochelle and on Route 10 toward Dover.)

various frequencies should be based upon measurements taken in the suburbs. The data from the New Rochelle and Dover series have been used as a basis for the points and the curve given in Figure 1. Each of the circle points shows the path loss at a given frequency relative to that at 150 mc for a particular location. Their spread indicates that the

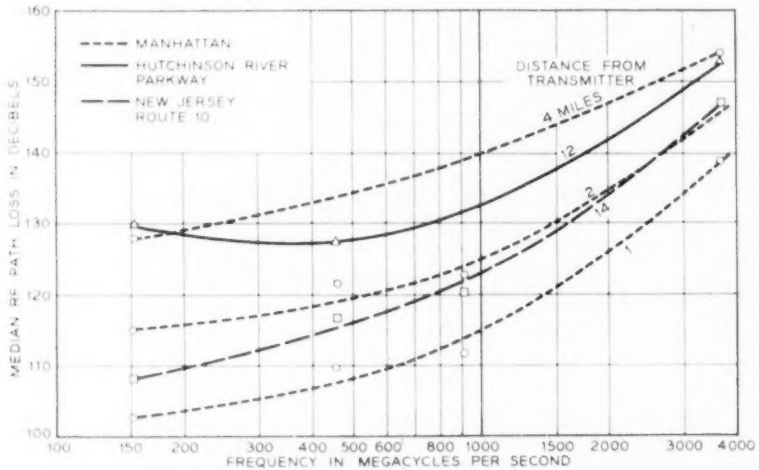


Fig. 10—RF path losses at locations for which 3700 mc measurements were made.

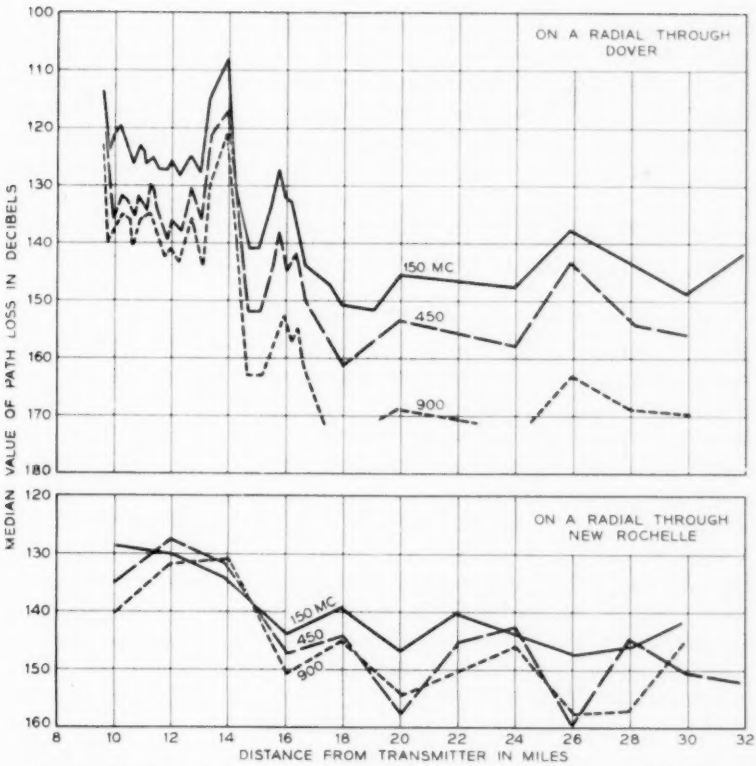


Fig. 11—Median r-f path losses along selected routes. (a) On a radial through Dover. (b) On a radial through New Rochelle.

comparison of frequencies is different at different locations. The "crosses" are the median values of these points, so placed that there are as many points above as below. The points for 3700 mc are taken from the data of Fig. 10. The crosses of Fig. 1 are considered to be the most reliable all-around comparison of propagation at the different frequencies.

RELATION OF SPEECH-NOISE RATIOS TO R-F SIGNAL POWER

Speech-to-noise ratios were measured at all of the test locations by the use of the level distribution recorder as described earlier. During the course of any given test the audio noise from the receiver varied considerably and these variations were recorded on the LDR. It was found by correlation between subjective observations of circuit merit and the

median value of noise that the latter is equivalent in noise effect to a steady random noise of the same value. In the FM receiver, the level of speech is essentially not affected by the strength of RF signal and so a measurement of the output noise is directly related to the speech-to-noise ratio. The speech-to-noise ratios given here are computed from noise measurements by assuming that speech of −14 vu level is applied to the system at a point where one milliwatt of 1,000 cycles tone would produce a 10-kc frequency deviation. The strength of the speech signal at the receiver output is expressed in the same units as are used for the noise.

As might have been expected the median speech-to-noise ratios correlate strongly with the amounts of r-f signal received at the various locations. This correlation has been evaluated in order that the most likely relationship between speech-to-noise ratio and received r-f signal may be known for the different frequencies. These are shown in Fig. 12, where each circle represents the median speech-to-noise value measured at one test location plotted against the median r-f signal received at that location. The solid lines have been drawn in to show the trend. The bending at the top of the curve is inconsequential. It only represents the limit imposed in the test setup by tube microphonic noise, vibrator noise, etc. The curves show, for example, that in order to produce a commercial grade of transmission, which requires a 12 db speech-to-noise ratio, the median r-f signal must be 122.5 db below one watt at 150 mc.

The data given in Fig. 12 pertain only to the suburban locations. Measurements in Manhattan have not been included, even though they indicate that larger signals are required, because the limit of system coverage is to be found in the suburbs. The data on the solid curves of Fig. 12 have been used to derive the curve of Fig. 2 which plots the value of r-f signal required at the mobile receiver for a commercial grade of transmission. The dotted curve of Fig. 2, which shows the median signal required in locations where noise picked up by the antenna is less than set noise, is based on the assumption of an 8 db noise figure for a practical 150-mc receiver, 11 db at 450 mc, and 12 db at 900 mc and higher.

Measurements have been made of the effect of noise picked up by the antenna at land receiver stations. These are expressed here in terms of the carrier strength required for just-commercial grade of transmission (12 db speech-to-noise ratio) as compared with the value required when there is no antenna noise and only receiver noise is present. These comparison measurements were made by injecting a steady carrier into the receiver with an antenna connected normally, and again with a dummy antenna connected. Although these tests were made with a steady rather

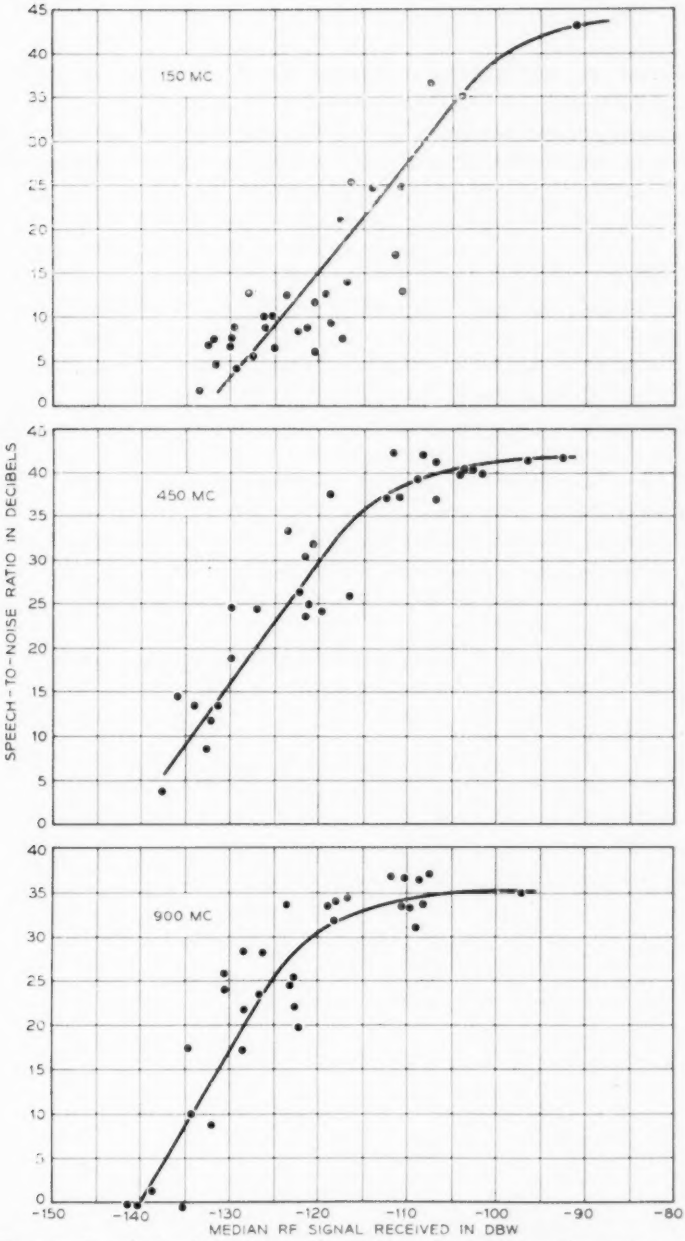


Fig. 12--Correlation between median values of speech-to-noise ratio and r-f signal strength in suburban locations. (Note: Each point represents the speech-to-noise ratio and the r-f signal received at one location. (a) 150 mc. (b) 450 mc. (c) 900 mc.

than randomly varying signal, it is felt that the comparative results will apply to the random signal case as well.

Tests were made at 150, 450 and 900 mc, at four locations of interest, and with dipole and 7 db gain antennas. Not all combinations were tested, but enough to permit some interesting comparisons. The locations tested were as follows:

A: On the Long Lines building, a 27-story building in downtown Manhattan.

B: On the Graybar-Varick building, a 16-story building in downtown Manhattan.

C: On the telephone building which houses the Melrose exchange, a 7-story building in the center of the Bronx.

D: On the 3-story telephone exchange building in Lynbrook, Long Island.

Table I describes the generally prevailing noise situation at these locations. Higher noise was encountered occasionally at some of the sites, due in at least one case to operation of elevators in the building. However, these occasions were so brief and infrequent that the general background of noise is considered to be a better value to use in estimating systems performance.

The trend toward lower site noise at higher frequencies, already noted for mobile installations, is seen to apply to land receivers as well.

TABLE I—R-F SIGNAL INPUT TO RECEIVER FOR 12 DB SPEECH-TO-NOISE RATIO (GIVEN IN DB ABOVE THAT NEEDED TO OVERRIDE SET NOISE*)

Station	Frequency	Antenna	
		Dipole	7-db Gain
A	150 mc	10	—
	450	1	0.5
	900	—	1.5
B	150	12	—
	450	4	—
	900	0	—
C	150	11	2.5
	450	1	1
	900	1	—
D	150	5	4
	450	0	0
	900	0	—

* Noise figures in the test receivers were 9, 12 and 12 db for 150, 450, and 900 mc, respectively.

These data bring out another interesting and significant fact. Where noise collected by a dipole antenna is discernible over set noise, the noise collected by the 7 db gain antenna at the same site is, surprisingly, less. This means that the gain antenna picks up *less* noise power than a dipole. Since it picks up 7 db more signal from a distant car, a gain antenna thus provides a double improvement in transmission at those sites for which ambient noise is controlling.

An explanation of this behavior may be surmised if it is assumed that the sources of noise are numerous and are scattered around at street level (motor vehicles, mostly). The overall noise received is a sum of contributions from all sources, weighted for distance and the receiving antenna pattern. A gain antenna of the type considered here tends to ignore the strong nearby noise sources because they are below the antenna beam. The sources, which are nearly enough in the beam to count, are also further away and are attenuated by distance.

The amount of data given in Table I does not seem sufficient to warrant stating a firm figure as to the amount of improvement obtainable from a gain antenna. However, substantial improvement at 150 mc is indicated, and this might have the effect of bringing the value of mobile transmitter power required at 150 mc down to the value required at 450 mc, assuming gain antennas in both cases.

ACKNOWLEDGMENTS

A number of people participated at one time or another in setting up and carrying through these tests. It is not possible to name them all, but the principal participants were R. L. Robbins, R. C. Shaw, W. Strack, D. K. White, and F. J. Henneberg. The program was supervised by D. Mitchell. The special radio equipment required was designed and furnished by W. E. Reichle and his group.

Common Control Telephone Switching Systems

By OSCAR MYERS

(Manuscript received August 1, 1952)

In the development of dial telephone switching systems two fundamentally different arrangements have been devised for controlling the operations of the switches. In one arrangement the switch at each successive stage is directly responsive to the digit that is being dialed. Systems using this method of operation are called direct dial control systems, an example being the step-by-step system as commonly used in the Bell System. In the other arrangement the dialed information is stored for a short time by centralized control equipment before being used in controlling the switching operations. Systems using the second arrangement are known as common control systems, examples of which are rotary, panel and crossbar. These two arrangements have different economic fields of use, the direct dial control being better suited for the smaller telephone exchanges and the common controls for the larger exchanges, especially those in metropolitan areas. A history of the evolution of these types of switching systems is presented, followed by a discussion of their comparative merits for various fields of use.

HISTORY

Invention of machines for switching telephone connections started shortly after the invention of the telephone. A forerunner of the step-by-step system, the Connolly and McTighe "girlless" telephone system,* was patented in 1879 and the first patent on the Strowger step-by-step system† was issued in 1891. The first commercial installation of automatic switching equipment was made at La Porte, Indiana, in 1892. This installation used step-by-step mechanisms.

In the early 1900's many telephone engineers regarded full automatic switching as uneconomical but technically feasible if restricted to single office exchanges with individual flat rate lines. They were, however, un-

* U. S. Patent 222,458—1879—Connolly and McTighe.
† U. S. Patent 447,918—1891—Almon B. Strowger.

certain about the future of this method of operation. It appeared to them that the greatest promise in the use of automatic apparatus was in distributing calls to manual "A" operators and in the elimination of the "B" operators. Consideration was being given to systems capable of operating on either a semi-mechanical or a full mechanical basis depending on whether the dial was located at the "A" board or at the subscriber's station. Development was also under way to provide arrangements for trunking calls between dial offices and to overcome the numerous weaknesses and deficiencies of existing dial systems.

The Strowger Company, the Bell System, and several other companies were planning or developing automatic and semiautomatic systems at that time. These included the full automatic, the network automatic, the automatic operator, and the semiautomatic. Short descriptions of some of them follow.

EARLY FULL AUTOMATIC SYSTEMS

The full automatic systems were mostly direct dial control. They included the Strowger, the Western Electric 100-line and 20-line, the Clark, the Faller* and the Lorimer systems.

The Strowger system of the middle 1890's provided 100-point two-digit selectors, one for each line. For each group of 100 lines the 100 outlets of each selector were multipled to the corresponding outlets of the other selectors serving the group. Each outlet of the group ran to a two-digit connector, each connector having access to 100 lines. Thus every group of 100 lines had 100 selectors and a maximum of 100 connectors and could reach 10,000 lines in a full office. Each group of connectors, up to the maximum of 100 connectors per group, had a multiple of 100 terminating lines. This was therefore a 4-digit single-office system theoretically of 10,000 lines capacity, requiring 1 selector and 1 connector per line. Subscribers in a given originating group of 100 lines had only one path to a particular terminating group of 100 lines. Since a selector was provided for each line, no dial tone was necessary. The switches used the familiar up and around motion. The exchanges of this type that were installed were small, the largest being in the order of 1000-line capacity. This type was followed by a new arrangement when automatic trunk selection was introduced. This provided multiple paths to each terminating group of 100 lines; the selector at this stage became a single-digit switch.

The Western Electric 100-line system could actually serve only 99

* U. S. Patent 686,892—Ernest A. Faller—Nov. 19, 1901.

lines. (The record does not disclose why one of the terminals of the system could not be assigned.) It used a rotary selector per line directly driven by a single train of pulses generated by a lever operated dial at the station. The selector had 100 points and the number of pulses sent corresponded to the number of the called line. The 20-line system was similar to the 100-line system.

The Clark system was a single motion rotary step-by-step system using 75-point switches which accommodated a maximum of 74 lines. (Here again there is no record as to why one terminal was not used for a line.) It did not provide a busy test. There were no relays in this system.

"AUTOMATIC OPERATOR" SYSTEMS

The Faller and the Lorimer systems were called "automatic operator" systems but they were actually versions of direct dial. The Faller system was apparently never used commercially, but the Lorimer system was.

The inventors of the Lorimer system had several objectives. One was to produce a system which could be installed in 100-line building blocks, called sections. As little as one section could be installed and operated alone. Additional sections in increments of 100-line capacity could be added as required up to the limit of 10,000 lines. Another object was to get good contacts and they therefore employed switches with heavy contacts like those used in power switches. The power needed to drive switches with such contacts led to the adoption of a common power drive for a number of switches instead of electromagnets individual to the switches. Still another aim was to provide a minimum of equipment on a per line basis and to provide equipment only to the extent required by traffic. Line relays were therefore omitted in early offices and the 100-line sections were divided into divisions, maximum 10 divisions per section, with arrangements for omitting divisions if not required by traffic.

The Lorimer system was a direct dial system operated from a pre-set calling device. It had a line finder stage, a selector stage and a connector stage. The calling device, wound up by a crank, had four settable levers, one for each digit, each of which grounded one terminal in its own set of ten terminals corresponding to the digit set up. The levers also operated a visual indicator. In the calling device there was also a switch driven over its terminals by a magnet-controlled escapement. Pulses were sent from the central office to control the escapement and the central office equipment was driven in synchronism with the station

switch until a grounded station terminal was found. The central office equipment was then stopped but the station switch continued stepping until the starting point for the next digit was reached. When the central office equipment was ready for the next digit the process was repeated until the called line was reached.

The Lorimer system has now disappeared from the scene in spite of a number of attractive features. The reasons for this disappearance are not clear from available records, but some reasonable conjectures can be made. For one thing, the pre-set calling device must have been expensive both in first cost and to maintain; it was also designed for a maximum of four digits and a re-design for more than four digits would have entailed substantial effort for developing both the calling device and the central office equipment. There is also some evidence to indicate that the system cost more than either step-by-step or panel.

THE NETWORK AUTOMATIC SYSTEM

The network automatic was a proposed form of semiautomatic in which the subscribers retained their manual instruments and were served by small unattended branch offices, each of which had a single group of trunks to a central operator office. On originating calls the branch offices acted as concentrators, automatically connecting calling lines to trunks to the central office where the operators were located and who asked for the called number as in straight manual practice. Called lines were reached through the branch offices by the operators at the central office who were provided with keysets to control the branch office equipment.

SEMI-AUTOMATIC SYSTEMS

There were several plans for other types of semi-automatic systems. Most of them contemplated replacing the "B" operator by a machine under control of the "A" operator. The plan of using machines under control of the "A" operators to replace the "B" operators was operated successfully in Saginaw, Mich. with Strowger apparatus. A similar plan was in operation in Los Angeles, and several groups of engineers studied improvements and variations.

STATUS IN 1905

The status of automatic switching by 1905 was this: there were several single office cities which had commercial installations of Strowger step-by-step equipment with severe limitations even for this field of use; a

number of Western Electric Company 100-line and 20-line automatics were in commercial service; a small amount of semi-automatic equipment was also in operation with the equipment under direct control of the "A" operator's dial; and planning and development work were under way to remove some of the limitations and extend the field of use of the automatic and semi-automatic systems.

The rotary dial was developed in 1896. However, many of the early systems did not use this type of dial. Various calling devices were used for a number of years. Among these were lever operated pre-set devices, keysets of several types, and dials with holes (in one case as many as 100) in which a peg could be inserted to act as a stop for an arm which was pulled around and allowed to restore. In all the early systems, regardless of the device used, the signals generated at the calling station directly controlled the selections.

RECOGNITION OF NEED FOR ACCESS TO LARGER TRUNK GROUPS

While mechanisms and circuits were being developed for direct dial control switching, work of a theoretical nature was going on which was to have an important effect on future designs. This work consisted of traffic probability studies and observations the outcome of which was the development of formulae and curves on the efficiency of trunk groups which influenced strongly the views of engineers as to the economical sizes of switches. G. T. Blood of the American Telephone and Telegraph Company in 1898 found that the binomial distribution closely fitted the observed data on the distribution of calls. The first comprehensive paper on the matter was one by M. C. Rorty in 1903, *Application of the Theory of Probability to Traffic Problems*. Curves accompanying his paper indicated that trunking efficiency improved with group size. Subsequent work by E. C. Molina in postulating that the grade of service experienced by a particular call applied to every call in the office and in developing the Poisson approximation to the binomial expansion formed the basis for trunking theory as used in the Bell System. Fig. 1 is a reproduction of three curves produced by Molina on July 6, 1908, showing the average load carried by various numbers of trunks for three probability conditions namely P.01, P.001 and P.0001 corresponding to an all trunks busy condition encountered by calls once in a hundred, once in a thousand, and one in ten thousand times respectively. From these curves it can be seen, for example, that ten trunks can carry a load averaging slightly over four calls with a probability of loss of P.01. Twenty trunks can carry an average of over eleven simultaneous calls

with the same P.01 loss but with an increase of efficiency of 15 per cent. The efficiency rises from 41 to 56 per cent.

EVOLUTION OF PRINCIPLE OF TRANSLATION

These studies had considerable effect on the trend of system design. For example, it appeared that grouping subscriber lines on the connectors in groups of more than 100 might result in some economy and that other economies were possible if the limitations imposed by decimal selections were avoided.

However, a new invention, namely translation, was required before systems could operate with large access switches and non-decimal selections. Translation is a mechanical rearrangement which permits conversion of the decimal information received from the dial to non-decimal forms for switch control and other purposes. When translation is made changeable by some means such as cross-connections, it is the basis of much of the flexibility of common-control systems. Translation was first proposed by E. C. Molina late in 1905. A patent application* for a *Translating and Selecting System* was filed on April 20, 1906.

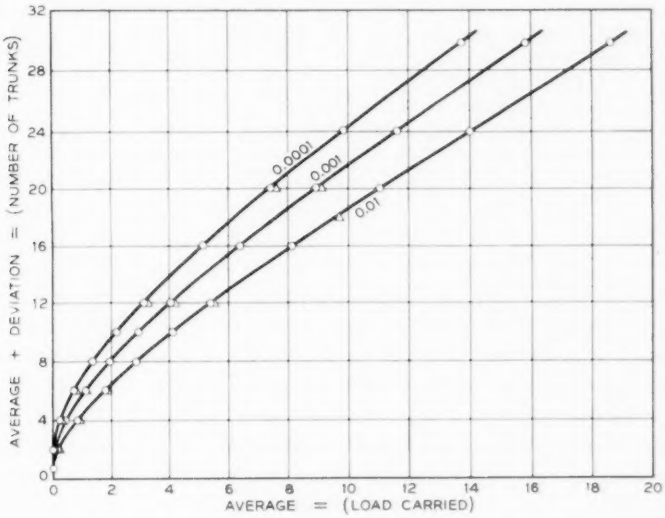


Fig. 2—Bypath system.

* Patent No. 1,083,456 issued to E. C. Molina, Jan. 6, 1916.

A necessary feature of systems employing translation of a series of digits such as an office code is digit storage. It was only a small step from the concepts of translation and digit storage to arrangements which provided these features in common circuits. Common controls with translation were first employed in the rotary system.

THE ROTARY AND PANEL SYSTEM DEVELOPMENTS

The rotary system was a full-fledged common-control system using register-senders to store the dialed information, to translate it to control the two-hundred point ten-level power-driven switches in selecting outgoing trunks from the originating office and in making line selections in the terminating office. The translation of the digits used for selecting trunks was changeable, but the translation of the numerical digits was fixed in permanent wiring of the register-senders.

In a search for less expensive cabling arrangements than those required by the rotary system, the panel bank employing punched metallic strips was developed. Each bank in the selectors of this system can accommodate 100 outlets with three wires per outlet, and five banks are stacked into a frame over which 60 power-driven selectors can hunt. For several years, starting in 1907, parallel development of the rotary and panel systems was carried on and desirable features of one were incorporated in the other. The panel system also has register-senders with changeable translation for selecting trunks and fixed translation for controlling selections in the terminating equipment. The major differences in the early designs of rotary and panel were due to the different access of the two systems and to differences in the methods of controlling the selectors. Both panel and rotary use revertive pulsing to control the selections. With revertive pulsing as the selectors progress they send back pulses which the sender counts. When a selector reaches the desired position, the sender stops it by opening the pulsing circuit. Both panel and rotary, like the Lorimer system, use a continuously operated power drive common to a number of switches because the increased size of switch which the greater access of these systems required, made a separate power drive economical.

The panel and rotary systems were originally designed for semi-mechanical operation with automatic distribution of calls to operators as a possible adjunct and with provision for full automatic operation if it proved desirable, by locating the dial or some other calling device at the subscriber's station rather than at the operator's position. This was a reasonable plan when development of these systems was started. Studies indicated that semi-mechanical systems could reduce the number of

operators required by an amount ranging from 30 to 50 per cent by eliminating the "B" operators and increasing the efficiency of the "A" operators. At that time, full automatic systems were subject to a number of shortcomings such as the complications and unreliability of the pulsing device at the subscriber's station, the need for a local battery at the station, and the lack of arrangements for party line and message rate service. Furthermore, there was considerable doubt as to the ability of the subscriber to dial with acceptable accuracy the six or seven numerical digits required in some of the multi-office exchanges.

There was an acute need for relief from the difficulties of manual operation after the start of World War I. Telephone growth was so rapid that it appeared for a time that the demand for new operators, particularly in the large cities, might outstrip the available supply. Competition from other industry for female help was also increasing. As more offices were added, the situation was further aggravated by the increasing complexity of operation. On account of the increasing number of trunked calls, the growing number of central offices, and the increasing amount of manual tandem operation, the quality of service was being degraded.

DEVELOPMENT OF A LARGE CITY NUMBERING PLAN

By 1916, the full automatic system (Strowger) had established a competitive position with manual for single-office cities, and both manual and full automatic offices were considered to be more economical than semi-mechanical for such cities. Because the number of dial pulls for a single office was four or less, little concern was felt about dialing accuracy.

For the multi-office cities it appeared that full mechanical operation would improve service and be more economical than either the semi-mechanical system or manual and would reduce the pressing need for operators. However, in spite of these factors urging the adoption of a dial system and even though automatic equipment was actually used in Los Angeles and Chicago in the first decade of the century, there was a reluctance to adopt full automatic operation in the very large multi-office cities because of the lack of a suitable numbering plan. A cumbersome plan was under consideration for handling dial traffic in these cities. This required the use of seven-digit numbers with the dial customers being called on to use arbitrary three-digit numerical codes for the office names. At the same time, the existing office names would be retained for use by the manual customers. Adoption of this dual arrangement would have required the provision of a cumbersome directory, but worse than that, it was felt that dialing seven numerical digits would be too

confusing to customers and that consequently there would be an excessive number of dialing errors. It was therefore planned to use semi-mechanical operation for cities like New York, retaining an operator between the customer and the machine. While this scheme did not save as many operators as the full mechanical method, it was believed necessary to have trained operators so that the customers would not be subjected to the complications of dialing. Under the proposed arrangement, the customer would pass the office name and number orally, and the operator would substitute the dial code for the office name and key or dial the code and number into the machine. Trial installations of the semi-mechanical panel system placed in service in the Waverly and Mulberry offices, Newark, N. J., in 1915 demonstrated that this method could provide reliable and improved telephone service under severe conditions.

However, in 1917 W. G. Blauvelt of the American Telephone and Telegraph Company proposed a numbering plan which would permit the customer to dial up to seven digits with acceptable accuracy and which would also be satisfactory for manual operation. This arrangement consisted of the use of one to three letters and four numbers. The first one, two or three letters of the office name were printed in bold type in the directory as an indication to dial customers that these were to be dialed ahead of the four numbers. Manual customers used the office name as before. Letters as well as numbers were placed on the dial plate in line with the finger holes of the dial. This proposal was immediately adopted and further Bell System development proceeded along the lines of full automatic operation. The Bell System planned to use the panel system in large cities not only because of the trunk efficiency which was possible with the use of the large panel switch, but also because trunking, being no longer under direct control of the dial in this system, was divorced from numbering. The panel system was also attractive because it had flexibility for growth and for contingencies such as the introduction of new types of service. These advantages would be provided by the common senders and translators of that system.

EARLY INSTALLATIONS OF COMMON CONTROL SYSTEMS

Early in 1918 tentative schedules were set up for 6-digit panel offices for Kansas City and Omaha and late that year a 7-digit office was recommended for the Pennsylvania office in New York City. When the Atlantic office in Omaha was placed in service on Dec. 10, 1921, it became the first commercial installation of a full automatic panel system.

Commercial installations of rotary equipment preceded the first com-

mercial panel offices. A semi-mechanical rotary system was installed in Landskrona, Sweden, in 1915 but remained in service for only a short time. A similar system was installed later in 1915 in Angiers, France. The first full mechanical rotary installation was at Darlington, England, in 1914. This system is still in service.

A common control system using Strowger switches, the director system, was developed in 1922. This development was prompted by the desire to provide automatic equipment in the London, England, multi-office exchange where the layout of the outside plant required considerable tandem trunking if a reasonably economical trunk network was to be achieved. All of the outside plant in London for the manual system was underground and it was required that this arrangement be retained when dial equipment was installed. This tended to fix the routes of telephone cables and to make it expensive to open new direct routes as new offices were opened. The trunking economies of tandems were extremely desirable under this condition and common controls with translation were necessary for a practical scheme capable of operating with the tandems. The director scheme, which in principle parallels the sender-translator scheme of the panel system, was designed to meet this situation. The director system was first placed in operation in Havana, Cuba, in 1924 and later in London in 1927.

EVOLUTION OF THE MARKER PRINCIPLE

In retrospect, it is obvious that the development thinking up to the early 1920's was limited by the belief that it was necessary to have the selectors do the testing for idle trunks even with common controls. This arrangement had been successfully used in the step-by-step system and it was natural to follow the same plan in the panel, rotary and director systems. Subsequent development of the common-control idea, starting with an experimental "coordinate" system in 1924, has resulted in marker systems in which the trunk testing is done by the markers.

The coordinate system derived its name from the method of operation of its switch, the process resembling the method of marking a point by the use of coordinates. The switch was essentially a large version of the crossbar switch and selected and held a set of crosspoints by the operation of horizontal and vertical members. Translation of the called office code, selection of a trunk, and operation of the switches to connect a transmission circuit to the trunk were functions of a new circuit, the marker, which the sender called into use for a fraction of a second after it had received the office code digits.

When the marker does the testing for idle trunks the trunk access from

a particular switch is no longer a limiting factor in the size of the trunk group. Once markers were invented it became possible to design systems using markers to do the trunk testing and any type of switch to do the connecting. When a trunk has been selected by the marker, the appropriate switches can be operated to connect to the marked terminal. The maximum size of trunk group need not be limited by the number of terminals on one switch. With a primary-secondary switch array groups much larger than those accessible on a single switch can be handled.

The coordinate system was not developed for commercial use. The first commercial marker system was PResident 2, a No. 1 crossbar office cut into service in Brooklyn, New York, in February, 1938. Improved crossbar systems have been developed since then including No. 5 crossbar and several types of toll crossbar systems

There is an interesting sidelight on the development of crossbar systems. The crossbar switch was invented by J. N. Reynolds of the Western Electric Company in 1913.* At that time proposed plans for using this switch assumed that it would be used as a line switch. The arrangements did not appear attractive and no serious attempt was made to develop a commercial system using the switch either as a line switch or as a selector. A number of years later an improved version of the crossbar switch was developed by the Swedish telephone administration. Their plans contemplated the use of the switch as a selector in a direct dial control system. In 1930 W. H. Matthies of Bell Telephone Laboratories visited Sweden and, impressed with the possibilities of the switch, ordered samples from Sweden after his return to the United States. Work was started to improve the switch and to develop a modern system around it. The crossbar switch, as previously mentioned, was a small version of the coordinate switch and the development of No. 1 crossbar was therefore started on a plan which was based on principles used in the coordinate system some of which had been successfully applied to the panel system with the adoption of the decoder in 1927.

TYPES OF COMMON CONTROL SYSTEMS

Four basic variations have been used in systems with common controls. These are (1) digit storage in common circuits on a decimal basis and control of switches by the stored digits without translation; (2) digit storage in the common circuits on a decimal basis, fixed translation and control of switches in a fixed pattern by the translated information; (3) a modification of the preceding plan in which the translation can readily

* U. S. Patent No. 1,131,734—J. N. Reynolds—issued March 16, 1915 and reissued December 26, 1916.

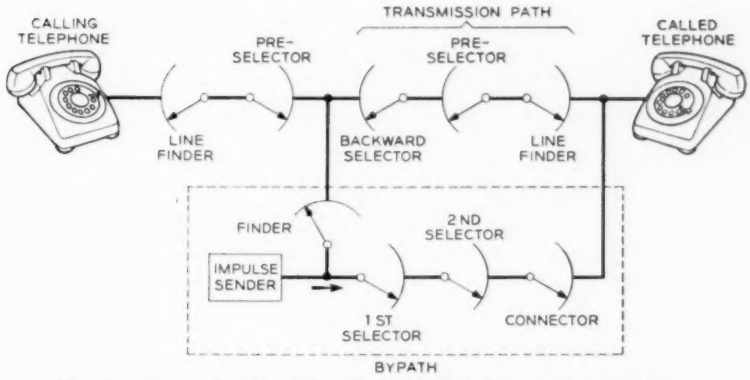


Fig. 1—Curves developed by E. C. Molina for trunk engineering.

be changed for any item of traffic; and (4) a still further variation where the function of hunting for an idle path is removed from the selectors and placed in new circuits called markers. Each variation resulted in improvements over preceding methods of operation.

The first plan is the simplest but also the least flexible. An advantage of this arrangement as well as of the other plans which also store the digits over step-by-step is that the interdigital time does not control the group size. By-path systems are examples of this method of operation. A system of this type is shown in Fig. 2. By-path systems use an auxiliary switch train that is under direct control of the dialed pulses to set up a connection. The talking circuit is then established over a parallel system of switches. The auxiliary train releases after the talking connection is set up and is available for use in setting up other connections. The Lorimer system avoided the penalties resulting from hunting during the interdigital interval by storing the digits at the station.

A further step in the direction of flexibility, but with added complication, can be taken by a *fixed* translation from a decimal to a non-decimal basis, i.e., a form of translation wherein a given decimal digit or a set of decimal digits is always changed into the same predetermined non-decimal equivalent. This permits the use of switches with less than ten groups of outlets thereby providing economies by permitting larger groups of outlets with a given size of switch.

A third variation with still greater flexibility than the first two, but also with greater complication, is a system with *changeable* translation. Changeable translation is achieved by providing some means such as cross-connections for readily changing the output pattern of the translators generally for sets of digits as, for example, for the called office

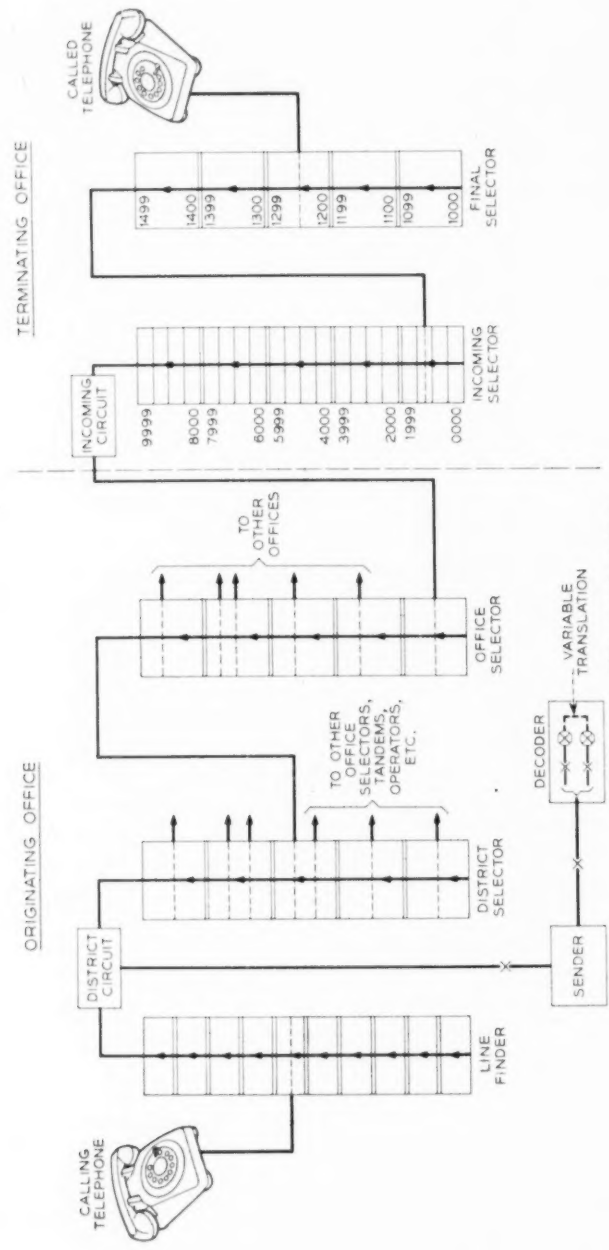


Fig. 3—The panel system.

codes. Changeable translation of office codes removes the limitation that the trunks for a given office designation must be located in a definite position on the switches which is the necessary result of fixed translation. Increased flexibility of numbering is now possible because office designation changes no longer require rearrangements of switch multiple. More economical arrays of switches are also possible because the switching plan can conform to traffic requirements without regard to numbering. Other advantages of translation—and as a practical matter, flexible translation—include the ability to operate with tandems, to operate with more than one type of outpulsing, and to operate with varying numbers of digits. The originating equipment of the panel system is an example of a system using changeable translation. This type of translation is also used for called line numbers as well as office codes in No. 1 and No. 5 crossbar thereby permitting these systems to shift lines for load balancing purposes without requiring numbering changes.

Finally, there is the most flexible but also the most complicated plan of all in which the selection of paths and trunks or lines is divorced from the selectors and placed in markers. In this plan the size of group is not limited by the number of terminals that a switch can hunt over in one sweep. No. 1 crossbar is an example of a system using the marker method of operation. In this system a switch generally has access to only ten trunks but on any one call a marker can test 160 trunks distributed over a number of switches.

Typical common control arrangements for systems using translation are shown in Fig. 3 for the panel system and in Fig. 4 for No. 1 crossbar.

The advantages noted are, in each case, the fundamental ones. Many others are inherent in common control and some will be brought out in further discussion.

A number of common control systems embodying the principles discussed have been designed. Rotary, panel and coordinate have been previously mentioned. Although the coordinate system never reached the commercial stage as a complete system, some of its features were adopted in the panel system starting in 1927 with the introduction of the decoder to replace the original three digit panel translator which used special panel selectors and pulse generating drums to do the translating job. This translator was limited in the digit combinations and number of three digit codes it could handle and also demanded a great deal of attention by the maintenance force. In place of the panel translators a small group of all-relay decoders, ranging from three to six, depending on traffic, was provided for each office. Senders were connected to decoders for about one-third of a second per call to obtain the information derived

from translation of the three office code digits. The connector for making the momentary connection of the large number of leads required between the senders and decoders presented new problems which were solved by the development of new relay preference and lockout circuits to permit as many simultaneous connections between senders and decoders as there were decoders and to permit an even distribution of calls to decoders. Decoder circuits were completely self-checking for trouble, provided for second trial in another decoder when trouble was discovered, and recorded troubles on a lamp bank trouble indicator.

In the early 1930's, encouraged by the success of decoders, the Bell System started development of the No. 1 crossbar system with markers in both originating and terminating equipments and with improved features over the coordinate system which it resembled in many respects. Self-checking circuits, second trials and trouble indicators which had proven highly successful in the decoder type panel system were important features of No. 1 crossbar. Automatic alternate routing and the ability to operate with non-consecutive PBX assignments were major new features introduced in this system for the first time.

The subsequently developed No. 5 crossbar system included a number of improvements, the chief of which from a common-control standpoint was the use of common markers for originating and terminating business and the use of the call back feature in setting up the connection. In this system the common equipment records the calling line identification as well as the called number, and after setting up to the called line or outgoing trunk, breaks down the connection to the common equipment

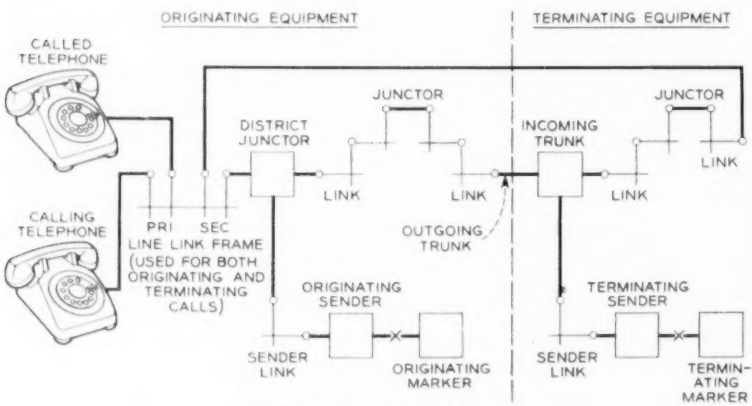


Fig. 4—No. 1 crossbar.

from the calling line and then re-establishes a connection back to the calling line.

Common controls have been employed by the Bell System in a number of systems in addition to those already mentioned. These include panel sender tandem, crossbar tandem, and No. 4, A4A and 4A toll crossbar.

COMPARISON OF COMMON CONTROL SYSTEMS AND DIRECT DIAL CONTROL SYSTEMS

Both direct dial control and common control systems have been developed to meet a wide range of situations for both large and small exchanges but, as previously noted, direct dial control systems have found their greatest field of use in the smaller exchanges and common control systems in the larger ones. The reasons for this can be brought out by a discussion of some of the features which have an important bearing on costs. These include the features affecting numbering plans, trunking arrangements, flexibility, quality of service, maintenance and engineering. A discussion of all the factors affecting costs will not be attempted. However, some of the more important ones will be covered.

RELATION BETWEEN TYPE OF SYSTEM AND NUMBERING PLANS

The requirements of a good numbering plan are well known. A good plan must be universal, i.e., must use the same number for reaching a called line regardless of the point of origin of the call in the area covered by the numbering plan, must permit dialing with acceptable accuracy, must permit directory listings that are readily understood by both dial and manual customers, and should use a minimum number of digits to reduce the labor of dialing. In small networks a satisfactory plan can be set up with almost any kind of system. However, especially in large networks, modern common control systems have outstanding advantages with respect to numbering.

These advantages of common controls are derived from the more flexible method of operation. Direct dial control systems use up the digits in the various stages of the switching operations whereas common control systems momentarily store them and can retransmit them. The result is that where direct dial control systems are used the numbering plan and the switching and trunking plans must conform whereas with common controls numbering, switching and trunking are not directly dependent on each other because the digits can be stored and translated. The effects of these differences on permissible latitude in numbering arrangements can be brought out by some examples.

Direct dial control systems cannot operate economically with a universal numbering plan in a network requiring any given call to have the possibility of completion over a variable number of links. The need for operating in this fashion arises when calls may be completed directly to the called office or via one or more tandem or toll systems. Numbering difficulties of a plan which attempts to use tandems with direct dial control systems can be illustrated by reference to Fig. 5. Assume that A, B, C represent three direct dial control type offices in a 6-digit numbering plan area and that these are connected by direct trunks between offices. Office B is designated ACademy (22 on the dial) and office C is designated BLue Hills (25 on the dial). Analysis of the trunk layout in this network indicates, let us say, that trunking economies can be made by establishing a tandem and that the direct route from A to C is no longer economical as compared to the route via the proposed tandem. The digits 25 must now select a route via tandem. However, if we use both digits for selecting the route to tandem we have none left for selecting the route to office C at the tandem office. Since this plan will not work, let us see what results if we assume that the tandem trunks are selected by means of the first digit. Now all calls starting with the code digit 2 at office A must be routed via tandem and even though economies call for a direct route to the ACademy office from A we are forced to use the uneconomical route through tandem for this office. Actually we must consider the economy of routing the traffic for all offices whose codes begin with a given digit via tandem, or routing it over direct trunks, or we must change the designation of one of the offices. We could, of course, adopt the undesirable expedient of using non-universal numbering, i.e., numbering that varied by points of origin, as, for example, by introducing extra digits on calls through tandem from A to C and omitting them on calls from B to C.

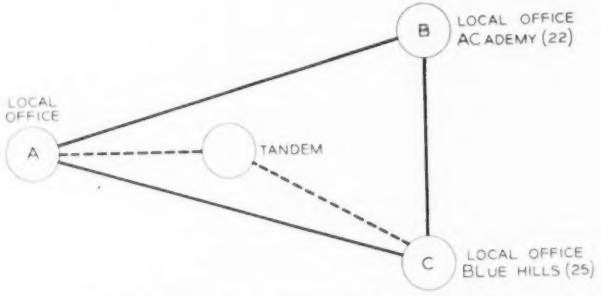


Fig. 5—Trunking scheme with a tandem office.

It is a situation such as has been described which has led to the practice, in some cases, of putting offices whose designations begin with the same first digit in the same building in step-by-step areas. This, of course, leads to restrictions.

Another alternative is to use selector repeaters. With these devices a "mitlaufer" action takes place in the local and tandem office selectors, i.e., both the local office selectors and the tandem office selectors follow the dial pulses until sufficient information is received to determine the route, whereupon the unneeded equipment is released. This equipment makes possible both the direct route to office B and the route via tandem to office C without an office designation change. However, selector repeaters are expensive and the cost of introducing them may be considerable. They also waste some trunk and equipment capacity because selector repeaters operate by seizing both local selectors and tandem trunks on every call. More often than not, perhaps, it would be cheaper to forego the trunk economy than to introduce the selector repeaters.

Now take the same network and assume common control equipment at all points. Prior to the introduction of the tandem the local offices translate the first two digits into information for selecting an outgoing trunk and then outpulse only the last four numerical digits directly to the called office. When the tandem is introduced, the translation at office A is changed to select a trunk to tandem on calls to BLue Hills and to tell the sender at A to spill ahead the code digits or equivalent information as well as the line number for these calls. For calls to ACademy the existing arrangement is retained. There is no special problem at tandem since the code for the called office, BLue Hills, is made available there. The translator at the tandem office tells the tandem sender to omit the office code digits in outpulsing to BLue Hills.

There is an essential difference in the coding between direct dial control and common control which is obscured by the use of the same codes in the examples. In the direct dial control case the codes are route codes (sometimes called group codes); that is, the digits directly correspond to the route through the switches and are expended in the switching operations. In the common control case they are destination codes and it is not necessary to have them conform to the route nor are they used up in the switching process. Only common control systems can operate with destination codes. Therefore common control systems are required where it is necessary to route calls to some offices by direct trunks and calls to other offices via tandems without numbering restrictions.

Another example of a numbering difficulty with direct dial control systems tracing back to the use of route codes, is illustrated by an

extreme example in Fig. 6. This figure shows a multi-switch route through four automatic intertoll switching systems, A, B, C, D, to a customer whose listed number is 2345 in the central office, MAin 2. MAin 2 is in numbering plan area 217, a different area from that of the calling office. Typical digit combinations are shown at each place for reaching the next place with direct dial control systems. On a call from the A toll center area to the number MA 2-2345, the originating toll operator must dial 16 digits, such as 059 076 097 157 2345. Calls starting at intermediate points or in other networks use different numbers depending on the route. (Note that the route codes start with 0 or 1 to distinguish them from local codes.) It is rather obvious that dialing such combinations is cumbersome and requires elaborate routing information at each toll center. Intertoll calls through direct dial control systems are therefore generally limited to being switched at one place along the route, with infrequent use of two switching points.

However, with common control systems the situation is quite different. The originating point need dial only the ten digits of the destination 217 MA 2-2345. At each point except the one preceding the called area the full complement of digits is sent ahead. At that point the area code is dropped. At the last point, D, which is assumed to have direct circuits to the called office, MA 2 is skipped and 2345 is sent ahead. If calling and called points had been in the same numbering plan area, only seven digits would have been required. Note that since destination codes are used all points outside the numbering plan area dial the same 10 digits to reach a given line and all points within dial the same seven digits.

While only a small proportion of toll calls require multi-switch connections of the type just described, connections such as these are nevertheless required for an economically feasible nationwide network in which all calls are dialed to completion, and this objective cannot be attained practically without systems operating with destination codes.

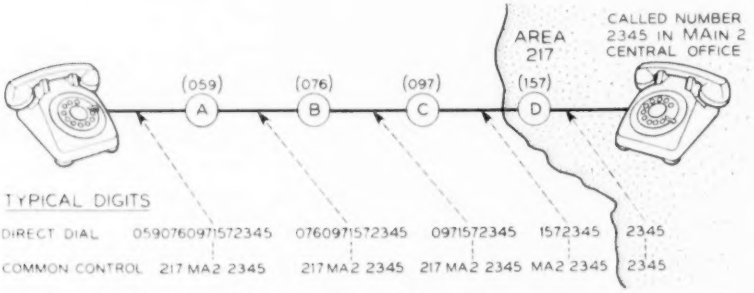


Fig. 6—Numbering with direct dial control and common control systems.

Also, as brought out later, destination codes are required in order to realize the important trunking economies of automatic alternate routing.

CODE CONVERSION

In passing, another feature of some common control systems, namely code conversion, can be brought out here because the illustration, Fig. 6, fits. Calls originating in a common control system can use office name codes (such as MA 2 for calls to the MAin 2 office) to reach destinations via step-by-step switching equipment where route codes (such as 157) are widely used. The translating equipment at the common control office can be arranged to substitute arbitrary digits for the office name code digits or in some cases to prefix arbitrary digits ahead of the called number. The arbitrary digits substituted or prefixed conform to the requirements of the office using route codes. In Fig. 6, office C when equipped with common controls could be arranged to convert MA 2 to 157, and therefore codes conforming to the nationwide numbering plan could be used for area 217 even though the calls were routed through step-by-step equipment.

RELATION BETWEEN TYPE OF SYSTEM AND TRUNKING ECONOMIES

The provision of a system which makes the most economical use of the trunk plant is important in any network but it is not as important in a small network as in a large one. Small networks can derive only small economies from arrangements which permit saving trunks. For example, in a single office network the trunks consist of wires running from originating to terminating equipment in the same building plus relatively cheap associated relay circuits. However, in a large toll network the trunks may include expensive repeaters, signaling equipments, carrier equipment and perhaps echo suppressors, as well as transmission channels running up to hundreds of miles in length and expensive toll relay circuits. For the larger networks there is therefore considerable urge to save as many trunks as possible. It is important therefore to operate these networks with switching plant that makes the most efficient use of the trunk plant by providing full access to groups, and to use an arrangement that permits the trunking economies of routes via tandems and of automatic alternate routing. These are features provided by common control systems and help explain why these systems are more attractive in the larger networks, both toll and local.

The cost of rearrangements for growth, new routes, load balancing and for restoring service under emergency conditions vary with the type

of system. Because of the flexibility of common controls such rearrangements are easier to make and usually cost less than in direct dial control systems. Also the frequency of rearrangements is greater in the larger places. Therefore this is another factor in favor of using common controls for those places.

SUPERIORITY OF COMMON CONTROL SYSTEMS WITH RESPECT TO SWITCH ACCESS

It has already been mentioned that the efficiency of trunks increases as the size of the group in which they are selected increases. Recognition of this fact early in the development of machine switching (about 1905) led to the invention of common controls. An ordinary step-by-step selector has access to only ten outlets on a level. Access to more than ten outlets can be obtained by providing graded multiple or by the use of rotary out-trunk switches,* or by combinations of these. Whenever it is necessary to employ graded multiple or rotary out-trunk switches, there is still some slight loss of efficiency as compared to full access.

In a system such as the panel system in which trunk hunting is a function of the selectors, the maximum number of trunks accessible to a call at any stage of selection is limited by the number of outlets accessible to the switch at that stage. A panel district or office selector, for example, can test a maximum of 90 trunks in a single group, 90 being the maximum number of terminals to which trunks can be assigned on a single panel bank, the remaining ten of the 100 terminals on a bank being reserved for overflow purposes. In the step-by-step system a corresponding limitation is avoided by a combination of graded multiple and rotary out-trunk switches with the penalty of a slight loss of efficiency. Marker systems avoid this limitation, also, by having the markers select trunks before they select the paths to the trunks. Crossbar systems with markers can readily test several hundred trunks for a given call. In some crossbar systems—No. 1, for example—trunks are tested in sub-groups of forty, therefore marker holding time is increased when there is more than one sub-group to be tested. This increase in marker holding time is largely avoided in systems like the toll crossbar systems by providing special testing arrangements in which a single indication per sub-group tells the marker which sub-group has one or more available trunks, whereupon the marker only tests the individual trunks of a sub-group in which it is assured that it can find an available trunk.

* A rotary out-trunk switch is arranged to hunt over a single group of outgoing trunks and to connect to an idle one. It is arranged for preselection and switches not in use will advance from busy trunks.

The maximum access of ten terminals on a level in ordinary step-by-step is not inherent in the system and might be overcome by a different switch design. A review of how a direct dial control system operates will help to clarify this point. At each switching stage, two actions take place. First, the switch follows the dial pulses until it reaches a group of outlets corresponding to the dialed digit. Then in the interval following this digit and before the pulses of the next digit arrive the switch hunts over the outlets for an idle path to reach the next stage. The number of paths from a switch level is therefore limited by the number of terminals the switch can hunt over in the interdigital interval. Assuming, for example, an interdigital interval of six-tenths of a second and a hunting speed of 100 terminals per second, 60 outlets could be provided. However, if such a high speed of hunting could be attained, and the 60 outlets were provided, 60 terminals would be required per group even for small ones which are in the majority. Hence such a switch would be wasteful of terminals. Direct dial control systems have generally employed switches with ten outlets per level although special arrangements such as twin levels have been employed to increase the number of outlets. A twin level switch provides terminals for two trunks at each rotary step and thus twenty trunks per level can be reached.

TRUNK ECONOMIES FROM TANDEM OPERATION WITH COMMON CONTROL SYSTEMS

An important factor in trunk economies is the ability to use tandems. The numbering difficulties that direct dial control systems have with tandems have already been discussed. Tandems permit major trunk economies on two scores. First, tandem routings take advantage of the efficiency which results from concentrating the smaller items of traffic and handling them over common trunk groups. Fig. 7 shows how this economy is attained. Ten offices completely interconnected by one-way trunks require 90 interoffice trunk groups. Ten offices interconnected only by way of tandem require only 20 groups. The groups by way of tandem are larger in size than the individual direct groups they replace and because of increased efficiency with group size fewer trunks are required.

There is a second possibility for an increase of efficiency, an example of which occurs when part of the offices are in business districts and part in residential districts. The peaks of trunked traffic from these different types of offices frequently occur at different hours, hence the trunks through tandem can be provided more economically for a given grade of

service than by an arrangement which must care for the peaks of each office separately. The non-coincidence of peaks of traffic of different types of offices permits economies both on trunks to tandem and trunks from tandem. For example, assume that a given office completes calls via tandem to some offices which have a morning busy hour and to others which have an evening busy hour. Then the group to tandem must provide capacity to handle the traffic for the busier hour of the two, but this capacity need care only for the peak traffic to part of the destinations. If individual direct groups had been provided instead of a common group to tandem, each group would have required capacity for its own peak, regardless of when it occurred. The common group to tandem therefore benefits by the noncoincidence of the peaks. A corresponding situation also occurs on trunks from tandem. Each group completes calls to a given destination from a number of originating offices whose peak hours may not coincide, and hence groups from tandem derive economies similar to those of the incoming groups to tandem.

Tandems are also required for alternate routing. Alternate routing is an arrangement to provide trunking economies by using a limited number of direct trunks for the traffic between two offices, and permitting the calls which do not find an available direct trunk to overflow to one or more tandems in succession. Because of the ability to load the direct circuits very heavily and yet provide good service by taking the overflow from and to a number of offices through a common tandem point, substantial economies are possible. Automatic alternate routing is practical only with common control systems. Common controls are needed to

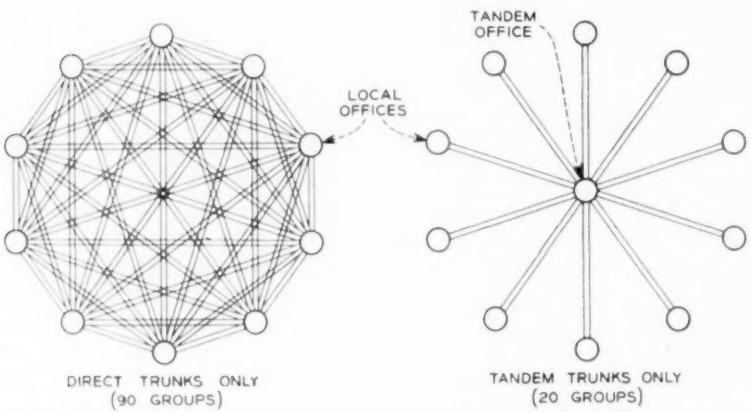


Fig. 7—Reduction of the number of trunk groups by the use of a tandem office.

provide the digit storage and digit spilling features in the office that does the alternate routing so that it can spill forward to the alternate route point the digits the latter requires.

Common controls have other advantages with respect to trunking which have already been covered in part. They also simplify the problems of assignment and load balancing as groups change in size or as new groups are added. An example of the difference in the methods of handling trunk growth in step-by step and crossbar is of interest. In step-by-step when groups grow beyond 10 trunks a grade must be introduced in the switch wiring, or trunks must be sub-grouped or rotary out trunk switches used. If further growth occurs, regrades must be made or rearrangements may be required in the sub-grouping or in the rotary out-trunk switches. In a crossbar system, however, in most cases added trunks are merely assigned to spare switch terminals which are left vacant for this purpose.

ROUTINGS FOR IRREGULAR CONDITIONS

Common controls are adapted to the efficient recognition and handling of irregular conditions such as permanent signals, vacant codes, and discontinued or temporarily intercepted lines.

Registers or senders detect line troubles which cause permanent signals or receivers off the hook by a timing circuit which waits for a short time for dialing to start. If the dialing does not start within the interval allowed the line is directed to a common group of permanent signal trunks which may appear before operators or at a test board. In No. 5 crossbar a trouble recorder card can be produced on which the location of the line in trouble is indicated. The step-by-step system indicates permanent signals by alarms to the maintenance force on a line group basis, and lines in trouble must be traced.

Vacant codes are detected by the translators, decoders and markers of common control systems and the calls are routed to a common trunk group which appears before operators or which returns "no such number tone." The corresponding arrangement in step-by-step requires connections from the switch multiple to operator or tone trunks.

In systems like No. 1 crossbar and No. 5 crossbar which have common controls in the terminating equipment, lines on which service has been discontinued or temporarily intercepted can be recognized by the markers and the calls rerouted to a common group of intercepting trunks. For example, temporary discontinuation of service is indicated by lifting a single cross-connection at the number group frame. In the step-by-step

system, however, one intercept trunk is commonly provided per 100 numbers and lines whose service is to be intercepted must be cross-connected to these trunks.

FURTHER ADVANTAGES OF COMMON CONTROL SYSTEMS ACCRUING FROM THEIR ABILITY TO OPERATE WITH TANDEMS

Some of the economies permitted by common control systems operating with tandems have been previously mentioned. Tandems are also useful because they provide centralized points at which special features can be concentrated with considerable saving.

For example, tandems are used for pulse conversion and for concentration of message charging equipment. Pulse conversion is needed when it becomes necessary to change from one type of pulsing to another, as, for example, on calls from a panel office to a step-by-step office. Panel can send out only revertive and panel call indicator pulses and step-by-step can receive only dial pulses. The two systems are therefore incompatible without special arrangements. The following are some of the plans which might be used for handling calls to step-by-step. First, all the panel senders could be modified to send out dial pulses. Second, spill senders could be provided at the outgoing trunks in the panel office or at the incoming trunks in the step-by-step office to receive, say, revertive pulses and convert them to dial pulses. Finally, if there is a tandem in the area, the tandem senders could be arranged (as they actually are) to accept revertive or panel call indicator pulses and send out dial pulses. The first two arrangements are usually more expensive than the last. Therefore, when pulse conversion is required it is generally done by routing calls via tandem.

To complete calls in the reverse direction, that is from step-by-step to panel, there is a requirement that is due to the use of the step-by-step system, namely that in cases where second dial tone is not employed the equipment at the called office or at an intervening tandem must be ready to accept the step-by-step pulses which are being dialed by the customer within a short time after the incoming trunk is seized. To meet this requirement, special high speed and costly link mechanisms are required to attach senders to incoming trunks or the incoming trunks must be arranged to record and store one or two digits. When calls are made between two systems both using senders, however, cheaper and slower link mechanisms can be employed because the calling senders are arranged to wait for a sender attached signal from the called office.

ADVANTAGES OF COMMON CONTROLS FOR AUTOMATIC RECORDING OF INFORMATION FOR CHARGING

The crossbar tandem system offers an economical method for making a record for charging purposes on multi-unit bulk billed calls called remote control of zone registration. At present this is limited to use with originating panel offices. The tandem is arranged to send back signals to the originating office for operating the customer's message register up to six times for the initial period on one call and also to operate it on overtime. Thus the application of extended customer dialing can be economically increased by applying this arrangement in places which cannot justify the registration arrangements available in the panel system itself which are economical only for a relatively heavy volume of this business. Local crossbar systems provide these features economically enough to obviate the need for tandem control of message registers for calls originating in the crossbar offices.

When tandem offices are required to control the equipment which records customers' charging data, they must be equipped with common controls if the arrangement is to be economical. The data includes the origin of the call—the particular trunk group incoming to tandem over which the call arrives—and the destination—the called office code. These elements must be analyzed and combined to determine the basis for the amount charged. Since elaborate equipment is required for these functions, economy must be attained by providing a minimum amount of equipment to do the job. This objective is accomplished by providing the required features in the common controls. In tandems arranged for remote control of zone registration, for example, the number of times the customer's message register is operated is determined partly by the choice of trunk group at the originating office and partly by the tandem markers.

In addition to remote control of zone registration, there are several other methods of determining and recording charging data which also require the use of common control equipment. These are automatic ticketing, automatic message accounting and coin zone dialing.

In automatic ticketing, which is used with step-by-step systems, calls which are to be ticketed are directed to outgoing trunks which select senders and other common equipment which determine the calling line number, reconstruct the called office code and store and outpulse the digits required for selections beyond the local office. The calling line number and the called office code are transmitted by the common equipment to the outgoing trunk which is equipped with a ticket printing

device which prints this information and other data required for charging. The tickets can be used for bulk bills as well as detail records since they can be summarized at the accounting center by manual methods for calls on which detail information is not required.

Automatic message accounting is used with crossbar systems both for bulk billing and detailed call records. With this system the data required for charging is perforated on paper tape by common central office equipment. The arrangement has been described in the technical literature* and will not be further described here.

Both the ticketing method and automatic message accounting require the collection of a large amount of data and the ability to do a complicated job in handling and recording this data. This demands elaborate and expensive equipment which is practical only when provided on a common basis so that it can be called into service for a short time and then restored to the common pool for other calls.

Direct dial control systems without common controls can only have message registers on the line and therefore can handle nothing but bulk billed calls. Furthermore because of the expense of arrangements for determining multiple unit charge data and for operating the message register more than once on a call, multiple operation of message registers on individual calls is not practical.

From coin stations in direct dial control systems the customer may dial calls only to offices within the local charge zone. However, in panel and crossbar areas the "coin zone dialing" arrangement is available to permit coin customers to dial beyond the local zone. With this plan calls are routed to a tandem office where completion is delayed until an operator can plug into the trunk to tandem and supervise the collection of the required coins. The amount to be collected is indicated by trunk lamps which appear in a switchboard multiple. Common controls enter into this scheme at the originating office to route the call to tandem and to determine the charge, and at the tandem office so that the digits can be stored while the call is held up prior to collection of the coins.

TYPES OF PULSING

Direct dial control systems are restricted to operation with dial pulses and are usually limited to pulsing speeds of about 10 pulses per second and about one digit per second. Dial pulsing has range limitations which can be overcome by the addition of pulse repeaters at appropriate points.

Common control systems store the digits in senders which can regen-

* *A.I.E.E. Transactions*, **69**, Part 1, pp. 255 to 268, 1950.

erate them in various types and combinations of types of pulsing. Types of outpulsing found today in various systems include revertive, panel call indicator, dial pulsing, dc key pulsing, and multi-frequency pulsing. Panel sender tandem and No. 4 toll can also send digital information ahead to operators by the call announcer method which uses voice announcements derived from recordings on film. Provision for receiving and sending several types of pulsing in one system makes it more flexible since it can then connect to a variety of equipments. Regenerating the pulses adds to the range without the need of adding pulse repeaters.

Some of the advantages which common control systems derive from the ability to operate with a modern type of pulsing can be brought out by a brief description of multi-frequency pulsing which is a relatively recent development. Digital information is transmitted over any facility capable of handling voice by sending spurts of alternating current which consist of pairs of frequencies in the voice range selected out of five frequencies. There are ten such pairs. At the receiving end a check is made to insure that exactly two frequencies are received for each digit. When only one or more than two frequencies per digit are detected the call is not set up but a reorder signal is returned to the originating end. In addition to the advantages of being capable of transmission over voice facilities, including repeaters and carrier systems, and of providing checks for accuracy, this type of pulsing can be transmitted at the rate of seven digits per second at present. Operators can be provided with keysets capable of sending MF pulses into either local or distant switching equipment with improved operating resulting from the higher speed and other advantages of MF pulsing.

It is quite feasible to add new types of pulsing to common control systems. Multi-frequency pulsing has only recently been added to crossbar tandem, for example, although it has been in use with other crossbar systems for some time. In this case it required the development of new senders capable of receiving and sending the MF pulses. The addition of these senders, even in existing offices, is not a difficult job.

IMPROVED STATION APPARATUS

The stations in most exchanges are provided with dials which operate at approximately 10 pulses per second. In step-by-step exchanges this pulsing speed is the maximum permitted by the capabilities of the switches. In panel and crossbar areas the common equipment is capable of operating with higher speed dial pulsing, and PBX and central office operators in these areas are usually given dials that operate at about 18 pulses per second.

Even fast dials are inefficient as compared to the push button keysets used by operators for key pulsing and it is obvious that subscriber sets with push buttons would be faster and more convenient than dials. Such sets were used at Media, Pa., on an experimental basis and have functioned in a highly satisfactory manner. Their introduction merely required the design and installation of registers to receive the pulses they generate. This was done with little difficulty or expense at the central office end. However, with ordinary step-by-step systems such devices are impractical because of the short interdigital interval they allow and because of the cost of adding the pulse receiving equipment in every selector and of providing translation to change the key pulses into a form to drive the switch.

CLASSES OF SERVICE

Differences in the handling of calls from non-coin, coin and PBX lines and differences in rate treatments require the recognition of classes of customers at the central office. In step-by-step separate groups of line finders are provided to permit segregation in classes and where routings for different classes vary, separate selector multiples are required for these routings. Class distinctions within a line finder group can be made by normal post springs and by marking a fourth conductor in the line circuit.

Common control systems permit the economical handling of many classes of service. The No. 5 crossbar, for example, is most flexible in this respect. As many as thirty classes of service can be handled in a single line link frame, including coin and non-coin. Special handling, reroutes and restrictions are mostly functions of the common controls and inefficiencies due to segregation of traffic in small groups of switching equipment are largely avoided.

DOUBLE CONNECTIONS

In systems such as panel and step-by-step in which selectors do the hunting, several selectors may be hunting over the same terminals simultaneously, and since there is an unguarded interval just after an idle terminal has been found before it is made busy by the release of the busy testing relay, double connections occur. Considerable effort and expense have been expended to reduce the probability of double connections in these systems. In systems which employ markers, on the other hand, the trunk testing schemes do not normally permit double connections to

occur. In most marker systems a lockout arrangement permits only one marker at a time to test trunks in a given group. There are cases where trunks are common to two offices and two markers are allowed to test trunks simultaneously. In these cases special circuit arrangements are provided at nominal expense to avoid double connections. Modern common control systems with markers are, therefore, free of double connections resulting from weaknesses of the system and they can occur only as a consequence of defects in circuits or apparatus.

THEORETICAL OFFICES

It is sometimes desirable to assign more than one office designation to customers in a single central office unit. A new unit may be planned for sometime in the future and if growth on the existing unit can be taken with a new office designation, then when this new office is placed in service it can be done without directory changes by transferring a block of lines from the old unit. Another occasion for assigning more than one designation to a single unit arises when customers served by the unit are in two rate zones, and service to lines in one of the rate zones must be restricted or extra charges collected. The lines served by an additional designation are called a theoretical office. Common control systems handle theoretical offices with little difficulty. In the first case mentioned the translating equipment in the originating offices recognizes that the physical office and theoretical office designations require identical treatment until the new unit is cut into service at which time translator cross-connection changes take care of the new routings. Where different rate treatments are involved, records for billing purposes depending on both the origin and destination of the call can be made by methods previously mentioned. In some cases where the billing data is determined at a tandem office and different treatments for the same destinations must be given to customers calling from one office, split trunk groups must be provided to tandem, one for each treatment.

In the step-by-step system, theoretical offices can be opened up by multipling two selector levels together. For example, if the physical office is designated 25 and it is desired to open a theoretical office, say 26, the 5 and 6 levels on the proper second selectors in the network can be strapped until the 26 office is changed to a physical office. At that time the levels are split and trunks to the new office are connected to the 6 levels of the second selectors. Restrictions in reaching blocks of numbers can be applied by splitting selector multiples and intercepting calls to restricted blocks from one of the splits.

ADAPTABILITY TO NEW SERVICE FEATURES

One of the major advantages of common controls, which has been covered in part but which deserves further emphasis, is their adaptability to new service features. Key sets and new dialing devices can be introduced at customers' stations and operator positions by readily feasible modifications of registers and senders. New pulsing schemes can also be introduced as they are developed as evidenced by the introduction of multi-frequency pulsing over the past few years. Nationwide customer dialing, now under development, can be readily introduced in existing common control systems by economical modifications without the use of either directing codes or second dial tone. Step-by-step systems require at least partial senderization to provide equivalent service. In short, the flexibility of common controls and the concentration of the control elements in a relatively few circuits makes the addition of new service features easier and more economical than in direct dial systems.

MAINTENANCE ASPECTS

Experience has shown that switches with a large amount of motion, especially those with brushes which wipe over bank terminals, tend to wear excessively and require considerable maintenance effort and even replacement, at times. On the other hand, switches with short motions and relay-like action require little maintenance and tend to have long life. Furthermore, the switches which employ wiping brushes mostly use base metal contacts, whereas relay-like switches can readily be equipped with precious metal contacts—and in most cases are so equipped—with the elimination of the transmission noise to which base metal contacts are subject. The crossbar switch is a relay type of switch with precious metal contacts and considerations such as those mentioned influenced its adoption. The advantages of relay type switches are not necessarily limited to common control systems since such switches have been used in direct dial control systems. The first use of the crossbar switch in Sweden was in a step-by-step system, for example. However, economical arrangements for using such switches in large systems require markers. This is because economy must be achieved by having more than one call occupy a switch at a time and marker control is necessary to attain this

Important maintenance advantages have been introduced in systems using decoders and markers. In this category are the self-checking features, second trials with changed order of preference, and trouble reporting features. In No. 5 crossbar the ability to report the location of a line

with a permanent signal by perforating a trouble recorder card has eliminated the need for tracing permanents.

A number of schemes are employed to detect troubles in markers and decoders and in circuits which connect to them. These include detectors for wrong sequences of operations, wrong combinations of relays, excessive current, false potential and lack of continuity. These are generally introduced at small cost since the circuits to which they are applied are small multipliers. However, some of them do a major job of testing since they reach out and test the numerous elements of the switching system to which markers have access. In this category are the tests of the crossbar linkages for opens, false grounds and double connections, tests of the switch crosspoints for continuity, tests of lines for false grounds, and for receivers off the hook on coin first coin lines.

To obtain clear trouble records, markers are designed with interlocked progress signals. This has made trouble analysis easier and has tended to improve design by eliminating relay races.

Starting with the panel system tests have also been introduced in senders for detecting open and reversed trunks. These tests have been of considerable help in maintaining outside plant and in detecting conditions that could lead to false charges.

DISADVANTAGES OF COMMON CONTROLS

Up to this point the stress has been mainly on the advantages of common controls. There are also some disadvantages. One of the major ones is the substantial getting started cost due to the necessity of providing a minimum amount of common equipment. This minimum is provided to maintain operation in case of trouble and during intervals when, for example, cross-connections require change because of changed or added routes. The minimum requirements establish economic barriers which tend to prohibit the economical use of common controls for small isolated systems.

Another disadvantage is the performance of common control systems under severe and protracted overloads. Experience with these systems indicates that although they compare quite favorably to direct dial control systems with respect to capability of handling moderate overloads, they are not able to handle severe overloads as well. In part this is a consequence of the fact that elements in common control systems are used at high efficiency and hence there is relatively less free equipment at full load for soaking up an overload than there is in systems that operate with smaller and less efficient groupings. Whenever the number

of calls presented to the system exceeds the capacity of the common control elements provided, the excess calls are delayed. The things which customers, operators and connecting switching machines do when they encounter delays tend to aggravate the overload. The reactions of operators and customers to delays can be illustrated by two examples.

The first is taken from the operation of a network of No. 4 toll crossbar systems when one of the No. 4's is heavily overloaded. Operators placing calls through the overloaded system encounter, let us say, an abnormal number of "no circuit" conditions in the outgoing trunks. This causes them to make additional attempts to get circuits. These additional attempts plus the excessive number of first attempts overload the markers. Sender holding time is then increased because of delays in connecting to the markers and this, added to the abnormal number of sender usages, results in a further shortage of senders. Operators trying to place calls through the system are therefore slowed down because of slow "sender attached" signals. (These are the signals which tell the operators that they can start keying or dialing.) Senders in connecting systems are also delayed waiting for senders to become idle in the overloaded office. The overload therefore tends to spread to all connecting systems.

However, it is possible to provide remedies which limit the reaction to the overloaded system. These remedies are arrangements to rapidly clear out senders waiting for senders ahead. Automatic alternate routing is also useful in routing traffic around overloaded systems.

The second example is taken from local systems. Here the reaction of customers to delays compounds the overload. A severe overload results in a shortage of senders, much as described above. A shortage of senders in a local system causes dial tone delays. There are always some customers who either do not listen for dial tone or who will not wait very long for it, and who start to dial before senders are attached to their lines. The result of such dialing is either a partial digits condition under which the sender waits for a considerable interval for a full complement of digits, or a wrong number when the first digit is clipped. The delays reduce sender capacity still further and the wrong numbers further increase the attempts. The load "snowballs" and the ability of the system to handle calls degenerates.

Here again arrangements are available to control the overload. These include features for blocking calls before they reach the senders and markers, and for returning paths busy signals with a minimum of common circuit holding time.

While there is, then, a somewhat greater capacity for overloads in step-by-step because of less efficient use of equipment, common control

systems do a good job of handling moderate overloads and, by provision of load control features, can operate satisfactorily even with severe overloads.

From a maintenance standpoint, a disadvantage of common controls is the relative complexity of the circuits. While this has introduced a training problem, maintenance forces have had no difficulty in acquiring the knowledge needed to do a competent maintenance job.

CONCLUSION

The full fledged common control systems exemplified by the crossbar local and toll systems have a number of important advantages over systems where the switches are driven directly by the customer's dial. The advantages arise largely from the ability to store digits, to translate them, use them flexibly for switching within the office, and transmit as many of them as desired to distant points for subsequent switching operations. The digits can be converted to others of different value whenever it is advantageous to do so. The inherent flexibility of common control equipment makes it possible to adopt any kind of numbering plan for a local area or a nationwide network that is best suited for the purpose without regard to the manner in which calls will be trunked from one point to another. Codes can be assigned at will to represent destinations and the best route for the call can always be taken. The best route may in some cases involve tandem operation or even a half-dozen switches in tandem. It may be the route selected as an alternate after previous trial of one or more other routes. A connection may be set up between offices of different types and over trunk groups requiring different forms of pulsing. These conditions may be met by common control equipment and the ability to meet such conditions makes it possible to provide cheap step-by-step equipment in places for which it is best suited, compensating for some of its deficiencies with common control equipment in other places.

With marker type common controls, trunk groups out of an office can be of any desired size regardless of the switch design. The individual crossbar switch, for example, gives access to only ten or twenty outlets as normally wired but full access single trunk groups of hundreds of trunks can be employed in some crossbar systems.

Schemes for recording billing data, aside from the relatively simple ones where metering equipment is associated with the customer's line and operated once per call, make use of common control equipment. This seems to be necessary where detail records must be made on individual calls for charging purposes.

As improvements in the art are made they can more readily be incorporated in common control systems than in step-by-step systems. For example, new subsets which may employ keys or other sending devices different from the dial can be accommodated by provision of proper facilities in senders and registers. Also, improved high speed pulsing arrangements can be easily incorporated in systems which do not require the switches themselves to be directly driven by pulses from the calling device.

BIBLIOGRAPHY

1. Bailey, W. J., "Lorimer automatic exchange at Hereford," *P. O. Elect. Engrs. J.*, **6,** Part 2, pp. 97–155, July, 1913.
2. Aitken, William, *Automatic telephone systems*, D. Van Nostrand Co., N. Y., 1924.
3. Miller, K. B., *Telephone theory and practice*, McGraw-Hill, N. Y., 1933.
4. Rorty, M. C., "How the theory of probability may be applied to telephone traffic," *Western Electrician*, **36,** p. 356, May 6, 1905, Abstract.
5. Craft, E. B., and others, "Machine switching telephone system for large metropolitan areas," Bell System Tech. J., **9,** pp. 266–274, Jan., 1934.
6. Bronson, F. M., "Tandem operation in the Bell System," Bell System Tech. J., **15,** pp. 380–404, 1936.
7. Scudder, F. J., and Reynolds, J. H., "Crossbar dial telephone switching system," Bell System Tech. J., **18,** pp. 76–118, Jan., 1939.
8. Abraham, L. G., and others, "Crossbar toll switching system," *A.I.E.E. Trans.*, **63,** pp. 302–309, 1944.
9. Friend, O. A., "Automatic ticketing of telephone calls," *A.I.E.E. Trans.*, **63,** pp. 81–88, 1944.
10. Smith, A. B., "The 'director' for automatic telephone switching systems," *A.I.E.E. Trans.*, **67,** pp. 611–619, 1948.
11. Meszar, J., "Fundamentals of the AMA system," *A.I.E.E. Trans.*, **67,** Part 1, pp. 255–269, 1950.

Mathematical Theory of Laminated Transmission Lines—Part II

By SAMUEL P. MORGAN, JR.

This part of the paper continues the analysis of the low-loss, broad-band, laminated transmission lines proposed by A. M. Clogston, and deals particularly with "Clogston 2" lines, in which the entire propagation space is filled with laminated material.

TABLE OF CONTENTS

VIII. PRINCIPAL MODE IN CLOGSTON 2 LINES WITH INFINITESIMALLY THIN LAMINAE

In Part I* of this paper we have set up a general mathematical framework for the analysis of Clogston-type laminated transmission lines and have applied it to Clogston 1 lines having laminated conductors, but with the total thickness of the laminations small compared to the overall dimensions of the line, so that most of the forward power flow takes place in the main dielectric. In Part II we shall consider Clogston 2 lines, which instead of containing a main dielectric have the propagation space entirely filled with laminations; and we shall also derive results, in Sections IX and X, for the general laminated transmission

* S. P. Morgan, Jr., *Bell System Tech. J.*, **31,** 883 (1952). Since the two parts of the paper are very closely related, the sections, equations, figures, and footnotes have been numbered consecutively throughout the whole paper. A table of symbols appears at the end of Part I.

line in which the relative fractions of space occupied by the main dielectric and the laminations are arbitrary.

A parallel-plane Clogston 2 line is shown schematically in Fig. 10. It consists of a stack of alternate layers of conducting and insulating material, whose total thickness is a. As before, the electrical constants of the conducting and insulating layers are denoted by μ_1, g_1 and ϵ_2, μ_2 respectively; and the fraction of conducting material in the stack is called θ. The stack is bounded at $y = \pm\frac{1}{2}a$ by sheaths whose normal surface impedance is $Z_n(\gamma)$, where γ is the longitudinal propagation constant of the mode under consideration.

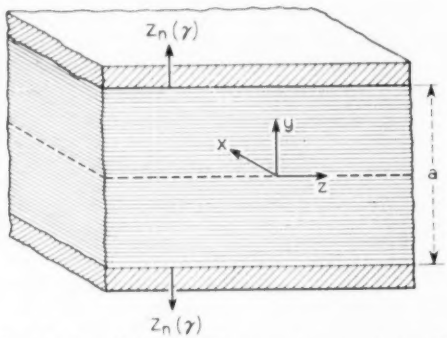


Fig. 10—Parallel-plane Clogston 2 transmission line.

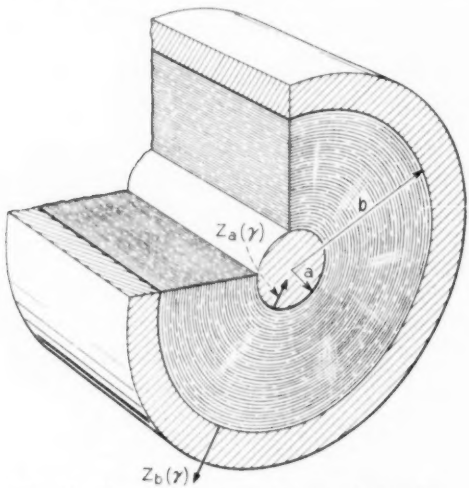


Fig. 11—Coaxial Clogston 2 transmission line.

The cross section of a coaxial Clogston 2 cable is shown schematically in Fig. 11. It consists of a laminated coaxial stack bounded internally by a cylindrical core of radius a, which may be equal to zero so far as the theoretical analysis is concerned, and externally by a cylindrical sheath of radius b. We denote the radial impedance looking into the core at $\rho = a$ by $Z_a(\gamma)$, and the radial impedance looking into the sheath at $\rho = b$ by $Z_b(\gamma)$.

In this section we shall assume the laminae to be infinitesimally thin, so that the stack may be regarded as a homogeneous, anisotropic medium, completely characterized by its average electrical constants. The case of finite lamina thickness will be treated in Section XI. We shall neglect dielectric and magnetic dissipation throughout, except in Section XIII.

For modes of the type which we consider, whose only field components are H_x, E_y, E_z in the plane line or H_ϕ, E_ρ, E_z in the coaxial line, the average electrical constants of the stack are given by equations (90) of Section III, namely,

$$\begin{aligned} \bar{\epsilon} &= \epsilon_2/(1-\theta), \\ \bar{\mu} &= \theta\mu_1 + (1-\theta)\mu_2, \\ \bar{g} &= \theta g_1 . \end{aligned} \tag{268}$$

As observed in Section III, Maxwell's equations for the average fields in such an artificial anisotropic medium take the form, for a plane stack,

$$\begin{aligned} \partial \bar{H}_x/\partial z &= i\omega\bar{\epsilon}\bar{E}_y, \\ \partial \bar{H}_x/\partial y &= -\bar{g}\bar{E}_z, \\ \partial \bar{E}_y/\partial z - \partial \bar{E}_z/\partial y &= i\omega\bar{\mu}\bar{H}_x ; \end{aligned} \tag{269}$$

while for a cylindrical stack,

$$\begin{aligned} \partial \bar{H}_\phi/\partial z &= -i\omega\bar{\epsilon}\bar{E}_\rho, \\ \partial(\rho\bar{H}_\phi)/\partial\rho &= \bar{g}\rho\bar{E}_z, \\ \partial \bar{E}_z/\partial\rho - \partial \bar{E}_\rho/\partial z &= i\omega\bar{\mu}\bar{H}_\phi . \end{aligned} \tag{270}$$

We wish to determine the modes which can propagate in the laminated medium when guided by plane or cylindrical impedance sheets. This problem was solved for a homogeneous, isotropic dielectric in Section II of Part I; and the method of solution is so similar for the anisotropic

medium that we shall omit details of the analysis and pass at once to the results.

In the parallel-plane line, the modes for which H_x is an even function of y about the center plane $y = 0$ have field components given by

$$\begin{aligned} H_x &= \operatorname{ch} \Gamma_\ell y \, e^{-\gamma z}, \\ \bar{E}_y &= -\frac{\gamma}{i\omega\bar{\epsilon}} \operatorname{ch} \Gamma_\ell y \, e^{-\gamma z}, \\ \bar{E}_z &= -K \operatorname{sh} \Gamma_\ell y \, e^{-\gamma z}, \end{aligned} \tag{271}$$

up to an arbitrary amplitude factor, where Γ_ℓ and K are defined, as in Section III, by

$$\Gamma_\ell = \left[\frac{i\bar{q}}{\omega\bar{\epsilon}} (\omega^2\bar{\mu}\bar{\epsilon} + \gamma^2)\right]^{\frac{1}{2}}, \tag{272}$$

$$K = \Gamma_\ell/\bar{q} = \left[\frac{i}{\omega\bar{\epsilon}\bar{q}} (\omega^2\bar{\mu}\bar{\epsilon} + \gamma^2)\right]^{\frac{1}{2}}. \tag{273}$$

Matching impedances at the boundaries $y = \pm\frac{1}{2}a$ leads to the condition

$$K \tanh \tfrac{1}{2}\Gamma_\ell a = -Z_n(\gamma). \tag{274}$$

In the odd case, the fields are given by

$$\begin{aligned} H_x &= \operatorname{sh} \Gamma_\ell y \, e^{-\gamma z}, \\ \bar{E}_y &= -\frac{\gamma}{i\omega\bar{\epsilon}} \operatorname{sh} \Gamma_\ell y \, e^{-\gamma z}, \\ \bar{E}_z &= -K \operatorname{ch} \Gamma_\ell y \, e^{-\gamma z}, \end{aligned} \tag{275}$$

up to an arbitrary amplitude factor, and the boundary condition becomes

$$K \coth \tfrac{1}{2}\Gamma_\ell a = -Z_n(\gamma). \tag{276}$$

General expressions for the field components in the coaxial line are

$$\begin{aligned} \bar{H}_\phi &= [AI_1(\Gamma_\ell\rho) + BK_1(\Gamma_\ell\rho)]e^{-\gamma z}, \\ \bar{E}_\rho &= \frac{\gamma}{i\omega\bar{\epsilon}} [AI_1(\Gamma_\ell\rho) + BK_1(\Gamma_\ell\rho)]e^{-\gamma z}, \\ \bar{E}_z &= K[AI_0(\Gamma_\ell\rho) - BK_0(\Gamma_\ell\rho)]e^{-\gamma z}, \end{aligned} \tag{277}$$

where A and B are arbitrary constants and Γ_ℓ and K are defined as before. The boundary conditions at $\rho = a$ and $\rho = b$ take the form

$$\begin{aligned} K\frac{AI_0(\Gamma_\ell a) - BK_0(\Gamma_\ell a)}{AI_1(\Gamma_\ell a) + BK_1(\Gamma_\ell a)} &= Z_a(\gamma), \\ K\frac{AI_0(\Gamma_\ell b) - BK_0(\Gamma_\ell b)}{AI_1(\Gamma_\ell b) + BK_1(\Gamma_\ell b)} &= -Z_b(\gamma), \end{aligned} \tag{278}$$

and these equations can be satisfied by values of A and B that are not both zero if and only if

$$\frac{KK_0(\Gamma_\ell a) + Z_a(\gamma)K_1(\Gamma_\ell a)}{KI_0(\Gamma_\ell a) - Z_a(\gamma)I_1(\Gamma_\ell a)} = \frac{KK_0(\Gamma_\ell b) - Z_b(\gamma)K_1(\Gamma_\ell b)}{KI_0(\Gamma_\ell b) + Z_b(\gamma)I_1(\Gamma_\ell b)}. \tag{279}$$

Now K is given in terms of Γ_ℓ by equation (273), while from equation (272) we have

$$\gamma^2 = -\omega^2\bar{\mu}\bar{\epsilon} - (i\omega\bar{\epsilon}/\bar{g})\Gamma_\ell^2. \tag{280}$$

Hence if the dependence of the boundary impedances on γ is known, equations (274) and (276) for the plane line and equation (279) for the coaxial line are transcendental relations from which in principle we may determine Γ_ℓ, and therefore γ, for each mode of the type that we are considering. If the value of Γ_ℓ for a particular mode satisfies the inequality

$$\frac{1}{8}\left|\frac{\Gamma_\ell^2}{\omega\bar{\mu}\bar{g}}\right|^2 \ll 1, \tag{281}$$

then on taking the square root of the right side of equation (280) by the binomial theorem, we find that the attenuation and phase constants of the given mode are approximately

$$\alpha = \operatorname{Re}\gamma = -\operatorname{Re}\frac{\Gamma_\ell^2}{2\bar{g}\sqrt{\bar{\mu}/\bar{\epsilon}}}, \tag{282}$$

$$\beta = \operatorname{Im}\gamma = \omega\sqrt{\bar{\mu}\bar{\epsilon}} - \operatorname{Im}\frac{\Gamma_\ell^2}{2\bar{g}\sqrt{\bar{\mu}/\bar{\epsilon}}}. \tag{283}$$

Throughout the rest of this section we shall consider only the lowest or principal mode. In a parallel-plane line the principal mode corresponds to the lowest root in Γ_ℓ (that is, the root having the smallest modulus) of equation (274), which may be written in the form

$$\tfrac{1}{2}\Gamma_\ell a \tanh \tfrac{1}{2}\Gamma_\ell a = -\tfrac{1}{2}\bar{g}aZ_a(\gamma). \tag{284}$$

We may express γ in terms of Γ_ℓ by equation (280), and so if $Z_n(\gamma)$ varies with γ in any reasonably simple way, or better yet if $Z_n(\gamma)$ is essentially independent of γ in the range of interest, equation (284) may be solved numerically for Γ_ℓ by successive approximations.

A numerical solution of equation (284) is, however, rarely necessary, since the right-hand side of the equation is just the ratio of the sheath impedance $Z_n(\gamma)$ to the resistance "per square", namely $1/(\frac{1}{2}\bar{g}a)$, of all the conducting layers in a stack of thickness $\frac{1}{2}a$ in parallel, and this ratio will almost always be large compared to unity. This is another way of saying that the total one-way conduction current in the stack is large compared to the sum of the conduction and displacement currents in either sheath. Even if the sheaths are infinitely thick metal plates of conductivity g_1, we have from equation (79) of Section III, since $\gamma \approx i\omega\sqrt{\bar{\mu}\bar{\epsilon}}$,

$$\tfrac{1}{2}\bar{g}aZ_n(\gamma) = \tfrac{1}{2}\theta g_1 a\eta_1 = (1 + i)\theta a/2\delta_1 , \tag{285}$$

and for most frequencies of interest the thickness $\frac{1}{2}\theta a$ of conducting material in half the stack will be several times the skin thickness δ_1 in the metal. If the medium outside the stack is free space, then $Z_n(\gamma)$ will be a few hundred ohms and a fortiori the right side of (284) will be large compared to unity. So long as the inequality

$$|\ \tfrac{1}{2}\bar{g}aZ_n(\gamma)\ | \gg 1 \tag{286}$$

is satisfied, the lowest root of (284) will be approximately

$$\Gamma_\ell = i\pi/a; \tag{287}$$

and so from (282) and (283) the attenuation and phase constants of a plane Clogston 2 line with infinitesimally thin laminae and high-impedance walls are

$$\alpha = \frac{\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}, \tag{288}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}. \tag{289}$$

To this approximation, there is neither amplitude nor phase distortion.

The principal mode in a coaxial Clogston 2 corresponds to the lowest root in Γ_ℓ of equation (279). To solve this equation numerically with finite boundary impedances $Z_a(\gamma)$ and $Z_b(\gamma)$, while possible in principle, would evidently be a major undertaking. We shall therefore assume throughout the present paper that the total conduction and displacement currents flowing in the core and the sheath are negligible compared

to the conduction currents in the laminated medium. This is equivalent to assuming that the boundary impedances $Z_a(\gamma)$ and $Z_b(\gamma)$ are effectively infinite, so that equation (279) reduces to the simple form

$$\frac{K_1(\Gamma_\ell a)}{I_1(\Gamma_\ell a)} = \frac{K_1(\Gamma_\ell b)}{I_1(\Gamma_\ell b)} . \tag{290}$$

Equation (290) may be converted to one involving ordinary Bessel and Neumann functions by the substitution

$$\Gamma_\ell^2 = -\chi^2, \qquad \Gamma_\ell = i\chi . \tag{291}$$

Then since

$$\begin{aligned} I_n(\Gamma_\ell \rho) &= i^n J_n(\chi\rho), \\ K_n(\Gamma_\ell \rho) &= \tfrac{1}{2}\pi i^{-(n+1)} [J_n(\chi\rho) - iN_n(\chi\rho)], \end{aligned} \tag{292}$$

the equation may easily be transformed into

$$J_1(\chi a)N_1(\chi b) - J_1(\chi b)N_1(\chi a) = 0. \tag{293}$$

For any given value of the ratio a/b, equation (293) has an infinite number of real roots in χ. The lowest root χ_1 has been tabulated[18] as a function of b/a, and may be written in the form

$$\chi_1 = \frac{\pi f_1(a/b)}{b - a}, \tag{294}$$

where $f_1(a/b)$ is a monotone decreasing function of a/b which is equal to 1.2197 when $a/b = 0$ and to 1 when $a/b = 1$. Hence the attenuation and phase constants of the principal mode in a coaxial Clogston 2 with infinitesimally thin laminae and high-impedance walls are

$$\alpha = \frac{\pi^2 f_1^2(a/b)}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\ \bar{g}(b - a)^2}, \tag{295}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}\,, \tag{296}$$

and again to this approximation there is neither amplitude nor phase distortion.

Comparing equations (288) and (295), we see that the attenuation constant of the principal mode in a coaxial Clogston 2 with infinitesimally thin laminae (that is, the low-frequency attenuation constant if the laminae are of finite thickness) is equal to the attenuation constant of

[18] E. Jahnke and F. Emde, *Tables of Functions*, fourth ed., Dover, New York, 1945, pp. 204–207. What we call $\pi f_1(a/b)$ is tabulated by Jahnke and Emde, p. 205, as $(k-1)x_1^{(1)}$, where $k = b/a$, while our $f_1(a/b)$ is plotted as $1 + \alpha$ on p. 207.

the principal mode in a plane Clogston 2 times the factor $f_1^2(a/b)$, provided that the thickness of the plane stack is equal to the thickness $b - a$ of the coaxial stack. The functions $f_1^2(a/b)$ and $f_1^2(a/b)/(1 - a/b)^2$ are plotted against a/b in Fig. 12. From the plots it is apparent that $f_1^2(a/b)$ decreases steadily from a value of 1.488 at $a/b = 0$ to 1 at $a/b = 1$, while $f_1^2(a/b)/(1 - a/b)^2$ increases steadily from 1.488 at $a/b = 0$ to infinity at $a/b = 1$. Therefore if the stack thickness $b - a$ is fixed, the attenuation constant will be smaller the greater is the mean radius of the stack; while if the outer radius b is fixed, the attenuation constant will be reduced by reducing the radius a of the inner core, and the lowest attenuation will be achieved when $a = 0$.

It should be noted that our expressions for the attenuation and phase constants of Clogston 2 lines cannot be valid down to the mathematical limit of zero frequency, since the inequality (281), on which we based the approximations (282) and (283) for α and β, will ultimately break down as the frequency approaches zero. A similar failure of the approximate expressions which we used for the attenuation and phase constants of Clogston 1 lines was pointed out in Section II of Part I. Here, as before, we shall limit the use of the term "low frequency" to frequencies still high enough so that the attenuation per radian is small and the approximate formulas (282) and (283) for α and β are valid.

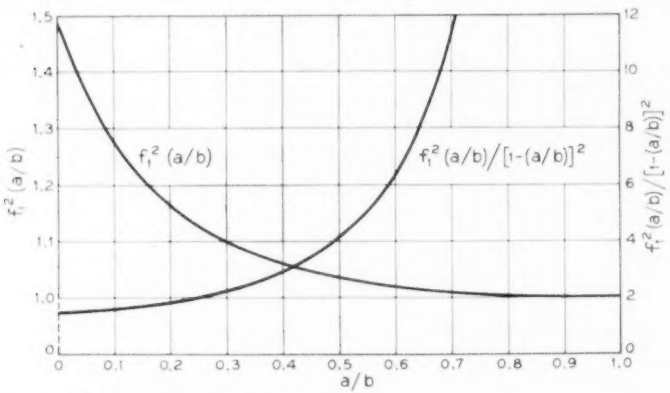


Fig. 12 –Curves related to the function

$$f_1^2(a/b) = (b - a)^2\chi_1^2/\pi^2,$$

where

$$J_1(\chi_1 a)N_1(\chi_1 b) - J_1(\chi_1 b)N_1(\chi_1 a) = 0.$$

Usually we shall be able to apply these formulas down to frequencies of a few kc·sec^{-1}.

The field components of the principal mode in a plane Clogston 2 with infinitesimally thin laminae and high-impedance boundaries at $y = \pm\frac{1}{2}a$ are given by equations (271), on substituting for Γ_ℓ from (287). We have, approximately,

$$H_z = H_0 \cos\frac{\pi y}{a}\, e^{-\gamma z},$$

$$\bar{E}_y = -\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 \cos\frac{\pi y}{a}\, e^{-\gamma z}, \tag{297}$$

$$E_z = \frac{\pi}{\bar{g}a}\, H_0 \sin\frac{\pi y}{a}\, e^{-\gamma z},$$

where H_0 is an arbitrary amplitude factor, and in the coefficient of the expression for $\bar{E}_y$ we have replaced γ by its approximate value $i\omega\sqrt{\bar{\mu}\bar{\epsilon}}$. The bars have been omitted from H_z and E_z since these field components are continuous at the boundaries of the laminae.

The potential difference between any two points in the same transverse plane is the integral of $-\bar{E}_y$ between the points. In particular, the total potential difference between the upper and lower sheaths is

$$V = -\int_{-\frac{1}{2}a}^{\frac{1}{2}a} \bar{E}_y\, dy = \frac{2a}{\pi}\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 e^{-\gamma z}. \tag{298}$$

The average value of the conduction current density $\bar{J}_z$ is

$$\bar{J}_z = \bar{g}E_z = \frac{\pi}{a}\, H_0 \sin\frac{\pi y}{a}\, e^{-\gamma z}, \tag{299}$$

and the current per unit width flowing in the positive z-direction in the upper half of the stack is

$$I = \int_0^{\frac{1}{2}a} \bar{J}_z\, dy = H_0 e^{-\gamma z}, \tag{300}$$

so that the ratio of voltage between the sheaths to total one-way current per unit width is

$$\frac{V}{I} = \frac{2a}{\pi}\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}. \tag{301}$$

The fields of the principal mode in a coaxial Clogston 2 with infinitesimally thin laminae and high-impedance boundaries are given by equa-

tions (277), which simplify somewhat if we write $i\chi_1$ for Γ_ℓ, replace the modified Bessel functions with ordinary Bessel functions according to (292), and remember that H_ϕ must vanish at the high-impedance boundaries. We then get, approximately,

$$\begin{aligned} H_\phi &= H_0[\ _1(\chi_1 b)J_1(\chi_1\rho) - J_1(\chi_1 b)N_1(\chi_1\rho)]e^{-\gamma z}, \\ \bar{E}_\rho &= \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0[N_1(\chi_1 b)J_1(\chi_1\rho) - J_1(\chi_1 b)N_1(\chi_1\rho)]e^{-\gamma z}, \\ E_z &= \frac{\chi_1}{\bar{g}}\, H_0[N_1(\chi_1 b)J_0(\chi_1\rho) - J_1(\chi_1 b)N_0(\chi_1\rho)]e^{-\gamma z}, \end{aligned} \tag{302}$$

where H_0 is an arbitrary amplitude factor.

The potential difference between any two points in the same transverse plane is the integral of $-\bar{E}_\rho$ between the points. Thus the total potential difference between the core and the outer sheath is

$$\begin{aligned} V &= -\int_a^b \bar{E}_\rho\, d\rho \\ &= \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, \frac{H_0}{\chi_1}[N_1(\chi_1 b)J_0(\chi_1\rho) - J_1(\chi_1 b)N_0(\chi_1\rho)]_a^b e^{-\gamma z} \\ &= \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, \frac{2H_0}{\pi\chi_1}\left[\frac{J_1(\chi_1 b)}{\chi_1 a J_1(\chi_1 a)} - \frac{1}{\chi_1 b}\right]e^{-\gamma z}, \end{aligned} \tag{303}$$

after some transformations using equation (293) and the well-known identity

$$N_0(x)J_1(x) - N_1(x)J_0(x) = 2/\pi x. \tag{304}$$

The average value of the conduction current density $\bar{J}_z$ is

$$\bar{J}_z = \bar{g}E_z = H_0\chi_1[N_1(\chi_1 b)J_0(\chi_1\rho) - J_1(\chi_1 b)N_0(\chi_1\rho)]e^{-\gamma z}. \tag{305}$$

The current reverses at $\rho = c$, where $a < c < b$ and c satisfies

$$N_1(\chi_1 b)J_0(\chi_1 c) - J_1(\chi_1 b)N_0(\chi_1 c) = 0; \tag{306}$$

hence the value of c may be found with the aid of a table of Bessel functions or from plotted curves.[19] The total one-way current in the outer part of the stack is

$$\begin{aligned} I &= 2\pi \int_c^b \bar{J}_z\rho\, d\rho \\ &= 2\pi c H_0[J_1(\chi_1 b)N_1(\chi_1 c) - N_1(\chi_1 b)J_1(\chi_1 c)]e^{-\gamma z} \\ &= -\frac{4H_0 J_1(\chi_1 b)}{\chi_1 J_0(\chi_1 c)}\, e^{-\gamma z}, \end{aligned} \tag{307}$$

[19] Reference 18, p. 208.

where in the last step we have made use of (304) and (306). The ratio of voltage across the stack to total one-way current is, from (303) and (307),

$$\frac{V}{I} = \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\,\frac{J_0(\chi_1 c)}{2\pi}\left[\frac{1}{\chi_1 b J_1(\chi_1 b)} - \frac{1}{\chi_1 a J_1(\chi_1 a)}\right]. \tag{308}$$

If there is no inner core, so that $a = 0$, the expressions which we have just derived become indeterminate forms, and it is simplest to make an independent calculation of the fields for this special case. The Neumann functions are now excluded because of their singularity at $\rho = 0$, and the condition (293) is replaced by

$$J_1(\chi b) = 0, \tag{309}$$

from which

$$\chi_1 = 3.8317/b. \tag{310}$$

The expressions for the fields are

$$\begin{aligned} H_\phi &= H_0 J_1(\chi_1\rho)e^{-\gamma z}, \\ \bar{E}_\rho &= \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 J_1(\chi_1\rho)e^{-\gamma z}, \\ E_z &= \frac{\chi_1}{\bar{g}} H_0 J_0(\chi_1\rho)e^{-\gamma z}, \end{aligned} \tag{311}$$

where H_0 is now a different arbitrary amplitude factor.

The total potential difference across the stack becomes, after putting in numerical values,

$$V = -\int_0^a \bar{E}_\rho\, d\rho = 0.3661\sqrt{\bar{\mu}/\bar{\epsilon}}\, H_0 b e^{-\gamma z}. \tag{312}$$

The conduction current density is

$$\bar{J}_z = \bar{g}E_z = \chi_1 H_0 J_0(\chi_1\rho)e^{-\gamma z}; \tag{313}$$

and $\bar{J}_z$ changes sign at

$$J_0(\chi_1 c) = 0, \qquad c = 2.4048/\chi_1 = 0.6276b. \tag{314}$$

The total one-way current is

$$I = 2\pi c H_0 J_1(\chi_1 c)e^{-\gamma z} = 2.047 H_0 b e^{-\gamma z}, \tag{315}$$

and the ratio of total voltage to total current is

$$V/I = 0.1788\sqrt{\bar{\mu}/\bar{\epsilon}}\,. \tag{316}$$

The fields of the principal mode in both plane and coaxial Clogston 2 lines will be plotted in the next section, when we shall also be able to show the fields in various transition structures between the extreme Clogston 1 and the complete Clogston 2.

As a numerical example, let us compare the attenuation constant of a conventional coaxial cable with that of a completely filled Clogston 2 cable of the same size. If a and b denote the radii of the inner and outer conductors of a conventional coaxial cable of optimum proportions ($b/a = 3.5911$), then at frequencies high enough to give a well-developed skin effect on both conductors, the attenuation constant is given by equation (151) of Section IV, namely

$$\alpha = \frac{1.796}{\eta_0 g_1 \delta_1 b}, \tag{317}$$

where η_0 is the intrinsic impedance of the main dielectric, which may be air. On the other hand, the attenuation constant of a Clogston 2 cable of outer radius b, with infinitesimally thin laminae and no inner core, is, from equations (282), (291), and (310),

$$\alpha = \frac{7.341}{\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g} b^2}. \tag{318}$$

It will be shown in the next section that for infinitesimally thin laminae whose permeabilities are all equal, the optimum value of θ is 2/3. Assuming no magnetic materials and setting $\theta = 2/3$, we find that the ratio of the attenuation constant α_c of the Clogston cable to the attenuation constant α_s of an *air-filled* standard coaxial cable of the same size, made of the same conducting material, is

$$\alpha_c/\alpha_s = 10.62 \sqrt{\epsilon_{2r}}\, \delta_1/b. \tag{319}$$

For copper conductors, δ_1 is given by equation (78) of Section III, and the crossover frequency above which the Clogston cable has a lower attenuation constant than the standard coaxial cable turns out to be

$$f_{\text{Mc}} = 763.5 \epsilon_{2r}/b^2_{\text{mils}}, \tag{320}$$

where frequency is measured in $\text{Mc} \cdot \text{sec}^{-1}$ and the radius of the cable in mils. We also note that at the crossover frequency the electrical radius of the inner conductor of the standard coaxial is $2.96 \sqrt{\epsilon_{2r}}\, \delta_1$, so that the use of equation (317) for α_s appears to be (barely) justified. Applying equation (320) to an ideal Clogston 2 cable of outer diameter 0.375 inches, excluding the sheath, with copper conductors, polyethylene

insulation, and no inner core, we have

$$b = 187.5 \text{ mils}, \qquad \epsilon_{2r} = 2.26, \tag{321}$$

and the crossover frequency is about 50 kc·sec^{-1}.

It must be emphasized that several factors which have not yet been taken into account will conspire to reduce the practical improvement in transmission that can be obtained with a Clogston 2 cable. As we have already seen for Clogston 1 lines in Part I, the effect of finite lamina thickness in a Clogston 2 will be to cause the attenuation constant to increase with increasing frequency, and ultimately to become higher than the attenuation constant of a conventional coaxial cable. Dissipation in the insulating layers may also contribute appreciably to the total loss at the upper end of the frequency band. Perhaps most important of all, the average electrical properties of the laminated medium must be held extremely uniform across the stack, or the field pattern of the principal mode will be distorted and its attenuation constant correspondingly increased. In later sections we shall discuss these effects, in order to estimate the stringency of the requirements on a physical Clogston cable if its factor of improvement over a conventional cable is to approximate closely to the theoretical limit given, for example, by equation (319).

IX. PARTIALLY FILLED CLOGSTON LINES. OPTIMUM PROPORTIONS FOR PRINCIPAL MODE

The distinction which has heretofore been made between Clogston 1 and Clogston 2 lines is rather artificial, inasmuch as both structures are limiting cases of the general Clogston-type line in which an arbitrary fraction of the total space is occupied by laminated material and the rest by an isotropic main dielectric. We shall now consider the modes which can propagate in a general partially filled line, restricting ourselves for simplicity to stacks of infinitesimally thin layers backed by high-impedance walls. Under these assumptions we first set up equations which must be satisfied by the propagation constants and the fields of all modes having only H_x, E_y, E_z or H_ϕ, E_ρ, E_z field components in a partially filled Clogston line, and then proceed to a study of the lowest or principal mode. We exhibit field plots for this mode at various stages of the transition between the extreme Clogston 1 and the complete Clogston 2 geometry, and investigate the conditions under which the attenuation constant passes through a minimum as the space occupied by the stacks is increased. This leads to the determination of certain optimum proportions for a line intended to transmit the principal mode.

In Section X we shall give a similar but briefer treatment of the various higher modes which can exist in partially or completely filled Clogston lines.

The notation for the parallel-plane line is established in Fig. 5 of Part I. The stacks are bounded externally by high-impedance sheaths at $y = \pm\frac{1}{2}a$, while the main dielectric is bounded by the planes $y = \pm\frac{1}{2}b = \pm(\frac{1}{2}a - s)$. No restrictions are placed on the relative thicknesses b and s of the main dielectric and the stacks. The average electrical constants of the stacks are $\bar{\epsilon}$, $\bar{\mu}$, and $\bar{g}$, while the electrical constants ϵ_0 and μ_0 of the main dielectric are assumed to satisfy Clogston's condition (102) but are otherwise arbitrary.

As in Section II, the modes may be divided into two classes, according to whether H_x is an even function or an odd function about the center plane $y = 0$. The normal surface impedance $Z(\gamma)$ looking into either stack may be obtained from equation (92) of Section III; if the impedance of the outer sheath is effectively infinite we have

$$Z(\gamma) = K \coth \Gamma_\ell s = (\Gamma_\ell/\bar{g}) \coth \Gamma_\ell s, \tag{322}$$

where

$$\Gamma_\ell = \left[\frac{i\bar{g}}{\omega\bar{\epsilon}}(\omega^2\bar{\mu}\bar{\epsilon} + \gamma^2)\right]^{\frac{1}{2}}. \tag{323}$$

Substituting for $Z(\gamma)$ into equations (11) and (13) of Section II, we find that the impedance-matching conditions become

$$\tanh \tfrac{1}{2}\kappa_0 b \tanh \Gamma_\ell s = -\frac{i\omega\epsilon_0}{\bar{g}}\frac{\Gamma_\ell}{\kappa_0}, \tag{324}$$

$$\coth \tfrac{1}{2}\kappa_0 b \tanh \Gamma_\ell s = -\frac{i\omega\epsilon_0}{\bar{g}}\frac{\Gamma_\ell}{\kappa_0}, \tag{325}$$

for the even and odd modes respectively, where

$$\kappa_0 = (\sigma_0^2 - \gamma^2)^{\frac{1}{2}} = (-\omega^2\mu_0\epsilon_0 - \gamma^2)^{\frac{1}{2}}. \tag{326}$$

From (323) we have

$$\gamma^2 = -\omega^2\bar{\mu}\bar{\epsilon} - (i\omega\bar{\epsilon}/\bar{g})\Gamma_\ell^2, \tag{327}$$

and so from (326),

$$\kappa_0^2 = -\omega^2(\mu_0\epsilon_0 - \bar{\mu}\bar{\epsilon}) + (i\omega\bar{\epsilon}/\bar{g})\Gamma_\ell^2. \tag{328}$$

If Clogston's condition is satisfied, namely

$$\mu_0\epsilon_0 = \bar{\mu}\bar{\epsilon}, \tag{329}$$

then

$$\kappa_0^2 = (i\omega\bar{\epsilon}/\bar{g})\Gamma_\ell^2, \qquad \kappa_0 = \sqrt{i\omega\bar{\epsilon}/\bar{g}}\ \Gamma_\ell, \tag{330}$$

and the equations for the even and odd modes become, respectively,

$$\tanh \tfrac{1}{2}\sqrt{i\omega\bar{\epsilon}/\bar{g}}\ \Gamma_\ell b \tanh \Gamma_\ell s = -\frac{\epsilon_0}{\bar{\epsilon}}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}} = -\frac{\bar{\mu}}{\mu_0}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}}, \tag{331}$$

$$\coth \tfrac{1}{2}\sqrt{i\omega\bar{\epsilon}/\bar{g}}\ \Gamma_\ell b \tanh \Gamma_\ell s = -\frac{\epsilon_0}{\bar{\epsilon}}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}} = -\frac{\bar{\mu}}{\mu_0}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}}. \tag{332}$$

For reference we shall now write down the field components of the various modes. The fields in the main dielectric are given by equations (8) and (12) of Section II, while the fields in the stacks may be obtained without difficulty if we recall that the tangential field components must be continuous at the inner boundary of each stack and that the tangential magnetic field must vanish at the high-impedance surface which forms the outer boundary of the stack.

Taking the even modes first, we have for the fields in the main dielectric,

$$\begin{aligned} H_x &= H_0 \operatorname{ch} \kappa_0 y\ e^{-\gamma z}, \\ E_y &= -H_0 \frac{\gamma}{i\omega\epsilon_0} \operatorname{ch} \kappa_0 y\ e^{-\gamma z}, \\ E_z &= -H_0 \frac{\kappa_0}{i\omega\epsilon_0} \operatorname{sh} \kappa_0 y\ e^{-\gamma z}, \end{aligned} \tag{333}$$

for $-\frac{1}{2}b \leqq y \leqq \frac{1}{2}b$, where H_0 is an arbitrary amplitude factor, γ and κ_0 are given in terms of Γ_ℓ by (327) and (330), and Γ_ℓ satisfies (331). The fields in the stacks are

$$\begin{aligned} H_x &= H_0 \frac{\operatorname{ch} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{sh} \Gamma_\ell(\tfrac{1}{2}a \mp y)\ e^{-\gamma z}, \\ \bar{E}_y &= -H_0 \frac{\gamma}{i\omega\bar{\epsilon}} \frac{\operatorname{ch} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{sh} \Gamma_\ell(\tfrac{1}{2}a \mp y)\ e^{-\gamma z}, \\ E_z &= \pm H_0 \frac{\Gamma_\ell}{\bar{g}} \frac{\operatorname{ch} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{ch} \Gamma_\ell(\tfrac{1}{2}a \mp y)\ e^{-\gamma z}, \end{aligned} \tag{334}$$

for $\frac{1}{2}b \leqq |\, y \,| \leqq \frac{1}{2}a$, where in case of ambiguous signs the upper sign is to be associated with the upper stack ($y > 0$) and the lower sign with the lower stack ($y < 0$). The continuity of E_z at $y = \pm\frac{1}{2}b$ is a consequence of equation (324) or (331).

For the odd modes, the fields in the main dielectric are

$$\begin{aligned} H_x &= H_0 \operatorname{sh} \kappa_0 y \, e^{-\gamma z}, \\ E_y &= -H_0 \frac{\gamma}{i\omega\epsilon_0} \operatorname{sh} \kappa_0 y \, e^{-\gamma z}, \\ E_z &= -H_0 \frac{\kappa_0}{i\omega\epsilon_0} \operatorname{ch} \kappa_0 y \, e^{-\gamma z}, \end{aligned} \tag{335}$$

for $-\frac{1}{2}b \leqq y \leqq \frac{1}{2}b$, where H_0 is again an arbitrary amplitude factor and γ and κ_0 are defined as before in terms of Γ_ℓ, which is now a root of (332). The fields in the stacks are

$$\begin{aligned} H_x &= \pm H_0 \frac{\operatorname{sh} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{sh} \Gamma_\ell(\tfrac{1}{2}a \mp y) \, e^{-\gamma z}, \\ E_y &= \mp H_0 \frac{\gamma}{i\omega\tilde{\epsilon}} \frac{\operatorname{sh} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{sh} \Gamma_\ell(\tfrac{1}{2}a \mp y) \, e^{-\gamma z}, \\ E_z &= +H_0 \frac{\Gamma_\ell}{\tilde{g}} \frac{\operatorname{sh} \frac{1}{2}\kappa_0 b}{\operatorname{sh} \Gamma_\ell s} \operatorname{ch} \Gamma_\ell(\tfrac{1}{2}a \mp y) \, e^{-\gamma z}, \end{aligned} \tag{336}$$

for $\frac{1}{2}b \leqq |y| \leqq \frac{1}{2}a$, where again the upper signs refer to the upper stack and the lower signs to the lower stack. The continuity of E_z at $y = \pm\frac{1}{2}b$ is now a consequence of equation (325) or (332).

The notation for the partially filled coaxial cable is shown in Fig. 6 of Part I, where as before we assume that the laminae are infinitesimally thin, the boundary impedances are effectively infinite, and the main dielectric satisfies Clogston's condition. The radius of the inner core is a and that of the outer sheath is b, while the stack thicknesses are s_1 and s_2 respectively; but no restrictions, other than obvious geometrical limitations, are placed on the relative values of a, b, s_1, and s_2. The inner and outer radii of the main dielectric are denoted by ρ_1 $(= a + s_1)$ and ρ_2 $(= b - s_2)$ respectively.

The boundary conditions at the surfaces of the main dielectric will be satisfied, as in Section II, by matching radial impedances at the stack-dielectric interfaces. If the impedance Z_a looking into the core at $\rho = a$ is effectively infinite, then the impedance looking into the inner stack at ρ_1 is given by equation (98) of Section III to be

$$Z_1 = \frac{\Gamma_\ell}{\tilde{g}} \frac{K_0(\Gamma_\ell\rho_1)I_1(\Gamma_\ell a) + K_1(\Gamma_\ell a)I_0(\Gamma_\ell\rho_1)}{K_1(\Gamma_\ell a)I_1(\Gamma_\ell\rho_1) - K_1(\Gamma_\ell\rho_1)I_1(\Gamma_\ell a)}. \tag{337}$$

Similarly, if the sheath impedance Z_b is infinite, then looking into the

outer stack at ρ_2 we have

$$Z_2 = \frac{\Gamma_\ell}{\tilde{g}} \frac{K_0(\Gamma_\ell \rho_2) I_1(\Gamma_\ell b) + K_1(\Gamma_\ell b) I_0(\Gamma_\ell \rho_2)}{K_1(\Gamma_\ell \rho_2) I_1(\Gamma_\ell b) - K_1(\Gamma_\ell b) I_1(\Gamma_\ell \rho_2)}. \tag{338}$$

The condition that the radial impedances shall be matched at the surfaces of the main dielectric is given by equation (38) of Section II, which takes the form

$$\frac{\kappa_0 K_0(\kappa_0 \rho_1) + i\omega\epsilon_0 Z_1 K_1(\kappa_0 \rho_1)}{\kappa_0 I_0(\kappa_0 \rho_1) - i\omega\epsilon_0 Z_1 I_1(\kappa_0 \rho_1)} = \frac{\kappa_0 K_0(\kappa_0 \rho_2) - i\omega\epsilon_0 Z_2 K_1(\kappa_0 \rho_2)}{\kappa_0 I_0(\kappa_0 \rho_2) + i\omega\epsilon_0 Z_2 I_1(\kappa_0 \rho_2)}, \tag{339}$$

where κ_0 is related to Γ_ℓ by equation (330). If we substitute the expressions (337) and (338) for Z_1 and Z_2 into (339), we have a single equation whose roots in Γ_ℓ correspond to all the circular transverse magnetic modes on the coaxial Clogston line. The propagation constant γ of each mode is given in terms of Γ_ℓ by equation (327).

Once the boundary conditions have been satisfied for a particular mode by a suitable determination of Γ_ℓ, it is a routine matter to obtain the field components for this mode. In the main dielectric the fields are of the form given by equations (33) of Section II. Hence for $\rho_1 \leqq \rho \leqq \rho_2$ we have

$$\begin{aligned} H_\phi &= [AI_1(\kappa_0\rho) + BK_1(\kappa_0\rho)]e^{-\gamma z}, \\ E_\rho &= \frac{\gamma}{i\omega\epsilon_0}[AI_1(\kappa_0\rho) + BK_1(\kappa_0\rho)]e^{-\gamma z}, \\ E_z &= \frac{\kappa_0}{i\omega\epsilon_0}[AI_0(\kappa_0\rho) - BK_0(\kappa_0\rho)]e^{-\gamma z}, \end{aligned} \tag{340}$$

where one of the constants A and B is arbitrary, but the ratio A/B must be taken equal to either side of equation (339). The fields in the stacks are of the form of equations (277) of Section VIII, where the constants are to be determined so that $H_\phi = 0$ at $\rho = a$ and $\rho = b$, and so that the tangential field components are continuous at ρ_1 and ρ_2. Imposing these conditions, we find that in the inner stack, for $a \leqq \rho \leqq \rho_1$,

$$\begin{aligned} H_\phi &= C[K_1(\Gamma_\ell a) I_1(\Gamma_\ell \rho) - I_1(\Gamma_\ell a) K_1(\Gamma_\ell \rho)]e^{-\gamma z}, \\ \tilde{E}_\rho &= \frac{\gamma}{i\omega\tilde{\epsilon}} C[K_1(\Gamma_\ell a) I_1(\Gamma_\ell \rho) - I_1(\Gamma_\ell a) K_1(\Gamma_\ell \rho)]e^{-\gamma z}, \\ E_z &= \frac{\Gamma_\ell}{\tilde{g}} C[K_1(\Gamma_\ell a) I_0(\Gamma_\ell \rho) + I_1(\Gamma_\ell a) K_0(\Gamma_\ell \rho)]e^{-\gamma z}, \end{aligned} \tag{341}$$

where

$$C = \frac{AI_1(\kappa_0\rho_1) + BK_1(\kappa_0\rho_1)}{K_1(\Gamma_\ell a)I_1(\Gamma_\ell\rho_1) - I_1(\Gamma_\ell a)K_1(\Gamma_\ell\rho_1)}. \tag{342}$$

In the outer stack, for $\rho_2 \leqq \rho \leqq b$,

$$\begin{aligned} H_\phi &= D[K_1(\Gamma_\ell b)I_1(\Gamma_\ell\rho) - I_1(\Gamma_\ell b)K_1(\Gamma_\ell\rho)]e^{-\gamma z}, \\ \bar{E}_\rho &= \frac{\gamma}{i\omega\bar{\epsilon}} D[K_1(\Gamma_\ell b)I_1(\Gamma_\ell\rho) - I_1(\Gamma_\ell b)K_1(\Gamma_\ell\rho)]e^{-\gamma z}, \\ E_z &= \frac{\Gamma_\ell}{\bar{g}} D[K_1(\Gamma_\ell b)I_0(\Gamma_\ell\rho) + I_1(\Gamma_\ell b)K_0(\Gamma_\ell\rho)]e^{-\gamma z}, \end{aligned} \tag{343}$$

where

$$D = \frac{AI_1(\kappa_0\rho_2) + BK_1(\kappa_0\rho_2)}{K_1(\Gamma_\ell b)I_1(\Gamma_\ell\rho_2) - I_1(\Gamma_\ell b)K_1(\Gamma_\ell\rho_2)}. \tag{344}$$

For the remainder of the present section we shall confine our attention to the principal mode. In the parallel-plane line this mode corresponds to the lowest root in Γ_ℓ of equation (331), that is,

$$\tanh \tfrac{1}{2}\sqrt{i\omega\bar{\epsilon}/\bar{g}}\ \Gamma_\ell b \tanh \Gamma_\ell s = -\frac{\bar{\mu}}{\mu_0}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}}. \tag{345}$$

We note that the right side of the equation is very small compared to unity, being of the order of the square root of the ratio of displacement current density in the insulators to conduction current density in the conductors, and also that the coefficient of Γ_ℓ in the first factor on the left will under all ordinary conditions be much smaller than the coefficient of Γ_ℓ in the second factor. Hence in seeking the lowest root we are justified in replacing the first hyperbolic tangent on the left side of (345) by its argument, so that the equation becomes

$$\Gamma_\ell s \tanh \Gamma_\ell s = -\frac{\bar{\mu}}{\mu_0}\frac{2s}{b}. \tag{346}$$

If we now let

$$\Gamma_\ell^2 = -\chi^2, \qquad \Gamma_\ell = i\chi, \tag{347}$$

we obtain

$$\chi s \tan \chi s = \frac{\bar{\mu}}{\mu_0}\frac{2s}{b}. \tag{348}$$

Since the right side of (348) is a positive real constant, the equation has

exactly one root in the interval $0 < \chi s \leqq \frac{1}{2}\pi$, which may most easily be found from a table[20] of the function $x \tan x$. If we call this root χ_1, equation (327) for the propagation constant γ becomes

$$\gamma^2 = -\omega^2\bar{\mu}\bar{\epsilon} + (i\omega\bar{\epsilon}\,\bar{g})\chi_1^2\,; \tag{349}$$

and on taking the square root by the binomial theorem we find for the attenuation and phase constants of the principal mode,

$$\alpha = \frac{\chi_1^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}\,, \tag{350}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}\,. \tag{351}$$

It is easy to verify that (350) reduces to the expressions previously obtained for the attenuation constants of Clogston 1 and Clogston 2 lines in the limiting cases $s \ll \frac{1}{2}a$ and $s = \frac{1}{2}a$ respectively. If $s \ll \frac{1}{2}a$, (348) gives

$$\chi_1^2 = \frac{2(\bar{\mu}/\mu_0)}{bs}\,, \tag{352}$$

so that from (350), on making use of Clogston's condition,

$$\alpha = \frac{(\bar{\mu}/\mu_0)}{\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}bs} = \frac{1}{\sqrt{\mu_0/\epsilon_0}\,\bar{g}bs}\,, \tag{353}$$

which agrees with equation (110) of Section IV. If $s = \frac{1}{2}a$, so that $b = 0$, then from (348),

$$\chi_1 = \tfrac{1}{2}\pi/s = \pi/a, \tag{354}$$

and (350) becomes

$$\alpha = \frac{\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}\,, \tag{355}$$

which is the same as equation (288) of the preceding section.

The general expressions (333) and (334) for the fields in a plane Clogston line with infinitesimally thin laminae and high-impedance walls simplify considerably for the principal mode, since κ_0 is so small that to a good approximation for $|\,y\,| \leqq \frac{1}{2}b$ we may replace sh $\kappa_0 y$ by $\kappa_0 y$ and ch $\kappa_0 y$ by unity. We then have, in the main dielectric,

[20] See for example Reference 18, Addenda, pp. 32–35.

$$H_x \approx H_0 e^{-\gamma z},$$

$$E_y \approx -\sqrt{\frac{\mu_0}{\epsilon_0}}\, H_0 e^{-\gamma z}, \tag{356}$$

$$E_z \approx \frac{\mu_0 \chi_1^2}{\bar{\mu}\bar{q}}\, H_0 y e^{-\gamma z},$$

for $-\frac{1}{2}b \leqq y \leqq \frac{1}{2}b$, where H_0 is an arbitrary amplitude factor. The fields in the stacks are

$$H_x \approx H_0 \frac{\sin \chi_1(\frac{1}{2}a \mp y)}{\sin \chi_1 s}\, e^{-\gamma z},$$

$$E_y \approx -\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 \frac{\sin \chi_1(\frac{1}{2}a \mp y)}{\sin \chi_1 s}\, e^{-\gamma z}, \tag{357}$$

$$E_z \approx \pm \frac{\chi_1}{\bar{q}}\, H_0 \frac{\cos \chi_1(\frac{1}{2}a \mp y)}{\sin \chi_1 s}\, e^{-\gamma z},$$

for $\frac{1}{2}b \leqq |y| \leqq \frac{1}{2}a$, where the upper signs refer to the upper stack and the lower signs to the lower stack. The potential and current distribu-

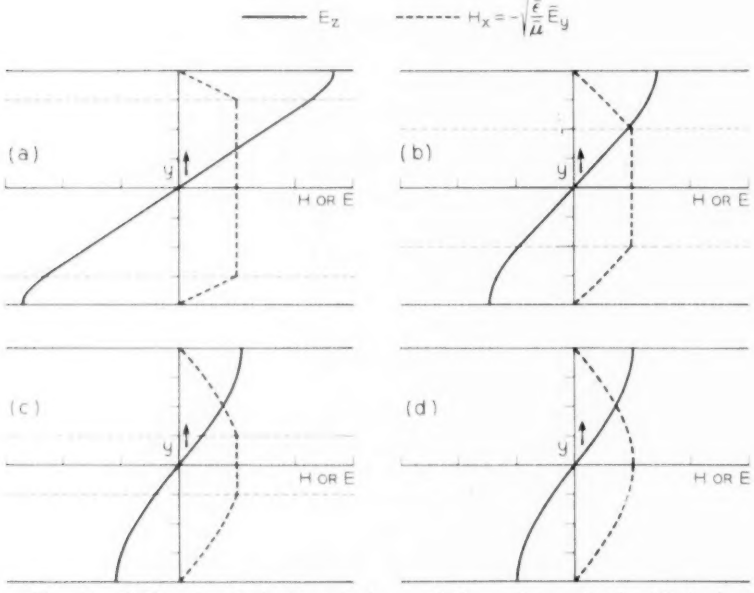


Fig. 13—Fields of principal mode in partially and completely filled plane Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

tions may easily be obtained, if desired, from the expressions for the field components.

As a numerical example we have plotted in Fig. 13 the fields of the principal mode for plane transmission lines in which the stacks fill respectively one-quarter, one-half, three-quarters, and all of the total available space. For simplicity we have taken $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$, and normalized the fields so that the total one-way current is the same in all four cases. The average current density is of course $\bar{g}E_z$ in the stacks and zero in the main dielectric. The first case approximates most nearly to the extreme Clogston 1 line discussed in Part I, while the last case is the complete Clogston 2, and the intermediate cases show the transition between these two structures. The following table gives $\chi_1 s$ as a function of the fraction $s/\frac{1}{2}a$ of the total space filled by the stacks, and also the quantity $(\chi_1 a/\pi)^2$, which is equal to the ratio of the attenuation constant of the line to the attenuation constant of the complete Clogston 2.

$s/\frac{1}{2}a$	$\chi_1 s \tan \chi_1 s$	$\chi_1 s$	$(\chi_1 a/\pi)^2$
$\frac{1}{4}$	$\frac{1}{3}$	0.5471	1.941
$\frac{1}{2}$	1	0.8603	1.200
$\frac{3}{4}$	3	1.1924	1.024
1	∞	1.5708	1.000

The principal mode in a coaxial Clogston cable corresponds to the lowest root in Γ_ℓ of equation (339). For the principal mode we are justified in assuming that in the main dielectric

$$| \kappa_0 \rho | \ll 1, \tag{358}$$

so that the Bessel functions occurring in equation (339) may be replaced by their approximate values for small argument. We thus find that, to a very good approximation, equation (339) reduces to the same form as equation (41) of Section II, namely

$$\frac{\kappa_0^2}{i\omega\epsilon_0} \log \frac{\rho_2}{\rho_1} = -\left(\frac{Z_1}{\rho_1} + \frac{Z_2}{\rho_2}\right), \tag{359}$$

where the stack impedances Z_1 and Z_2 are given by (337) and (338). If as before we let

$$\Gamma_\ell^2 = -\chi^2, \qquad \Gamma_\ell = i\chi, \tag{360}$$

then from (330) κ_0^2 becomes

$$\kappa_0^2 = -(i\omega\bar{\epsilon}/\bar{g})\chi^2; \tag{361}$$

and replacing the modified Bessel functions in (337) and (338) with ordinary Bessel functions according to equations (292) of Section VIII, we obtain

$$Z_1 = \frac{\chi}{\bar{g}} \frac{J_1(\chi a)N_0(\chi\rho_1) - N_1(\chi a)J_0(\chi\rho_1)}{J_1(\chi a)N_1(\chi\rho_1) - N_1(\chi a)J_1(\chi\rho_1)}, \tag{362}$$

$$Z_2 = \frac{\chi}{\bar{g}} \frac{J_1(\chi b)N_0(\chi\rho_2) - N_1(\chi b)J_0(\chi\rho_2)}{J_1(\chi\rho_2)N_1(\chi b) - N_1(\chi\rho_2)J_1(\chi b)}. \tag{363}$$

Substituting (361), (362), and (363) into (359) and setting $\bar{\epsilon}/\epsilon_0 = \mu_0/\bar{\mu}$, we get after a little rearrangement,

$$\begin{aligned} &\frac{1}{\chi\rho_1} \frac{J_1(\chi a)N_0(\chi\rho_1) - N_1(\chi a)J_0(\chi\rho_1)}{J_1(\chi a)N_1(\chi\rho_1) - N_1(\chi a)J_1(\chi\rho_1)} \\ &\quad + \frac{1}{\chi\rho_2} \frac{J_1(\chi b)N_0(\chi\rho_2) - N_1(\chi b)J_0(\chi\rho_2)}{J_1(\chi\rho_2)N_1(\chi b) - N_1(\chi\rho_2)J_1(\chi b)} = \frac{\mu_0}{\bar{\mu}} \log \frac{\rho_2}{\rho_1}. \end{aligned} \tag{364}$$

If χ_1 is the smallest positive root of equation (364), then the attenuation and phase constants of the principal mode are

$$\alpha = \frac{\chi_1^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}, \tag{365}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}. \tag{366}$$

These expressions for α and β are of exactly the same form as equations (295) and (296) for a complete Clogston 2, except that χ_1 is now determined from equation (364) instead of equation (293). It is easy to show that (364) reduces to (293) if there is no main dielectric, that is, if $\rho_1 = \rho_2$.

For any given values of the four ratios a/b, ρ_1/b, ρ_2/b, and $\mu_0/\bar{\mu}$, equation (364) may be solved for $\chi_1 b$ by numerical or graphical methods. However if we wish to examine many cases, so as to investigate the effects of varying some or all of the parameters, a more efficient procedure for finding χ_1 is needed. Such a method is provided by the observation that in spite of the complicated appearance of equation (364), it is really just the equation which determines the eigenvalues in a rather simple two-point boundary-value problem, which is well adapted to solution on a differential analyzer. We digress briefly to formulate this problem.

The differential equations for the fields in the main dielectric can be put in the form of equations (67) of Section III, namely

$$\begin{aligned} d(-\rho H_\phi)/d\rho &= -i\omega\epsilon_0\rho E_z, \\ dE_z/d\rho &= -(\kappa_0^2/i\omega\epsilon_0\rho)(-\rho H_\phi), \end{aligned} \tag{367}$$

where the propagation factor $e^{-\gamma z + i\omega t}$ has been suppressed. On the other hand, equations (270) for the fields in a stack of infinitesimally thin laminae yield

$$\begin{aligned} d(-\rho H_\phi)/d\rho &= -\bar{g}\rho E_z, \\ dE_z/d\rho &= -(\Gamma_\ell^2/\bar{g}\rho)(-\rho H_\phi), \end{aligned} \tag{368}$$

where Γ_ℓ is given by (323). If we neglect the displacement current in the main dielectric compared with the conduction current in the stacks, replace Γ_ℓ by $i\chi$, express κ_0 in terms of χ by (361), and write $\mu_0/\bar{\mu}$ for $\bar{\epsilon}/\epsilon_0$, we obtain the following equations for the fields in the coaxial Clogston line:

For $a \leqq \rho \leqq \rho_1$,

$$\begin{aligned} d(-\rho H_\phi)/d\rho &= -\rho(\bar{g}E_z), \\ d(\bar{g}E_z)/d\rho &= (\chi^2/\rho)(-\rho H_\phi); \end{aligned} \tag{369i}$$

while for $\rho_1 \leqq \rho \leqq \rho_2$,

$$\begin{aligned} d(-\rho H_\phi)/d\rho &= 0, \\ d(\bar{g}E_z)/d\rho &= (\mu_0\chi_c^2/\bar{\mu}\rho)(-\rho H_\phi); \end{aligned} \tag{369ii}$$

and for $\rho_2 \leqq \rho \leqq b$,

$$\begin{aligned} d(-\rho H_\phi)/d\rho &= -\rho(\bar{g}E_z), \\ d(\bar{g}E_z)/d\rho &= (\chi^2/\rho)(-\rho H_\phi). \end{aligned} \tag{369iii}$$

The quantities $-\rho H_\phi$ and $\bar{g}E_z$ must be continuous at ρ_1 amd ρ_2; and the two-point boundary condition at the infinite-impedance surfaces $\rho = a$ and $\rho = b$, namely

$$-aH_\phi(a) = -bH_\phi(b) = 0, \tag{370}$$

determines a sequence of eigenvalues $\chi_1^2, \chi_2^2, \chi_3^2, \cdots$, of which the lowest corresponds to the principal mode. It is a routine matter to integrate equations (369) in terms of Bessel functions and logarithms, and to show that the continuity and boundary conditions lead exactly to equation (364).

If equations (369) are set up on a differential analyzer with adjustable values of χ^2, a/b, ρ_1/b, ρ_2/b, and $\mu_0/\bar{\mu}$, it is a simple procedure to make a few runs with different choices of χ^2, and so to locate the approximate

value of χ_1^2 which satisfies the boundary conditions for the given values of the other parameters. If additional accuracy is wanted, it is then not too difficult to refine this approximate value by desk computation. The results of quite a number of exploratory calculations which were made on the Laboratories' general purpose analog computer will be shown later in this section.

The fields of the principal mode in the main dielectric of a Clogston cable are given approximately by equations (46) of Section II, namely

$$\begin{aligned} H_\phi &\approx \frac{I}{2\pi\rho}\, e^{-\gamma z}, \\ E_\rho &\approx \sqrt{\frac{\mu_0}{\epsilon_0}}\,\frac{I}{2\pi\rho}\, e^{-\gamma z}, \\ E_z &\approx \frac{I}{2\pi \log\,(\rho_2/\rho_1)}\left[\frac{Z_1}{\rho_1}\log\frac{\rho_2}{\rho}+\frac{Z_2}{\rho_2}\log\frac{\rho_1}{\rho}\right] e^{-\gamma z}, \end{aligned} \tag{371}$$

for $\rho_1 \leqq \rho \leqq \rho_2$, where I is an amplitude factor equal to the total current flowing in the positive z-direction in the inner stack, and Z_1 and Z_2 are given by writing χ_1 for χ in (362) and (363) respectively. The fields in the inner stack are

$$\begin{aligned} H_\phi &\approx \frac{I}{2\pi\rho_1}\,\frac{N_1(\chi_1 a)J_1(\chi_1\rho) - J_1(\chi_1 a)N_1(\chi_1\rho)}{N_1(\chi_1 a)J_1(\chi_1\rho_1) - J_1(\chi_1 a)N_1(\chi_1\rho_1)}\, e^{-\gamma z}, \\ \bar{E}_\rho &\approx \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\,\frac{I}{2\pi\rho_1}\,\frac{N_1(\chi_1 a)J_1(\chi_1\rho) - J_1(\chi_1 a)N_1(\chi_1\rho)}{N_1(\chi_1 a)J_1(\chi_1\rho_1) - J_1(\chi_1 a)N_1(\chi_1\rho_1)}\, e^{-\gamma z}, \\ E_z &\approx \frac{\chi_1}{\bar{g}}\,\frac{I}{2\pi\rho_1}\,\frac{N_1(\chi_1 a)J_0(\chi_1\rho) - J_1(\chi_1 a)N_0(\chi_1\rho)}{N_1(\chi_1 a)J_1(\chi_1\rho_1) - J_1(\chi_1 a)N_1(\chi_1\rho_1)}\, e^{-\gamma z}, \end{aligned} \tag{372}$$

for $a \leqq \rho \leqq \rho_1$; while in the outer stack we have

$$\begin{aligned} H_\phi &\approx \frac{I}{2\pi\rho_2}\,\frac{N_1(\chi_1 b)J_1(\chi_1\rho) - J_1(\chi_1 b)N_1(\chi_1\rho)}{N_1(\chi_1 b)J_1(\chi_1\rho_2) - J_1(\chi_1 b)N_1(\chi_1\rho_2)}\, e^{-\gamma z}, \\ \bar{E}_\rho &\approx \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\,\frac{I}{2\pi\rho_2}\,\frac{N_1(\chi_1 b)J_1(\chi_1\rho) - J_1(\chi_1 b)N_1(\chi_1\rho)}{N_1(\chi_1 b)J_1(\chi_1\rho_2) - J_1(\chi_1 b)N_1(\chi_1\rho_2)}\, e^{-\gamma z}, \\ E_z &\approx \frac{\chi_1}{\bar{g}}\,\frac{I}{2\pi\rho_2}\,\frac{N_1(\chi_1 b)J_0(\chi_1\rho) - J_1(\chi_1 b)N_0(\chi_1\rho)}{N_1(\chi_1 b)J_1(\chi_1\rho_2) - J_1(\chi_1 b)N_1(\chi_1\rho_2)}\, e^{-\gamma z}, \end{aligned} \tag{373}$$

for $\rho_2 \leqq \rho \leqq b$. The potential and current distributions may be calculated in the usual way from the fields.

As numerical examples we have plotted in Fig. 14 the fields of the principal mode in two Clogston coaxial cables. Fig. 14(a) shows a

cable in which $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$, and with the dimensions $a = 0.084b$, $\rho_1 = 0.415b$, and $\rho_2 = 0.831b$, these proportions having been found optimum, as discussed below, for a cable without magnetic loading in which the total thickness of both stacks is arbitrarily chosen equal to $\frac{1}{2}b$. Fig. 14(b) shows the fields of a complete Clogston 2 with no inner core, the scale being chosen so that the total one-way current is the same in both cases. The attenuation constant of the first cable is 1.234 times that of the second one. Fig. 14 may be compared with Fig. 13 for the plane geometry, whence it should be possible to visualize approximately the fields of the principal mode in other coaxial structures representing various stages of the transition between the extreme Clogston 1 and the complete Clogston 2 cable.

Now let us consider a Clogston line with infinitesimally thin laminae, having fixed external dimensions and containing only materials with given electrical constants. We may pose two questions: (1) Supposing that for some practical reason the total available thickness of laminated material is also fixed, how should this material be divided between the two stacks to minimize the attenuation constant of the line? (2) Assuming that the total thickness of laminations in the line is at our disposal, what is the optimum amount of laminated material from the standpoint of minimizing the attenuation, and how should this material be distributed in the optimum case?

For plane transmission lines the first question is trivial; the stacks should always be of equal thickness. In a coaxial cable, if the filling ratio $(s_1 + s_2)/b$ is given, the proportions of the cable are completely determined when we specify, for example, the relative radius a/b of the core and the relative thickness $s_1/(s_1 + s_2)$ of the inner stack. The opti-

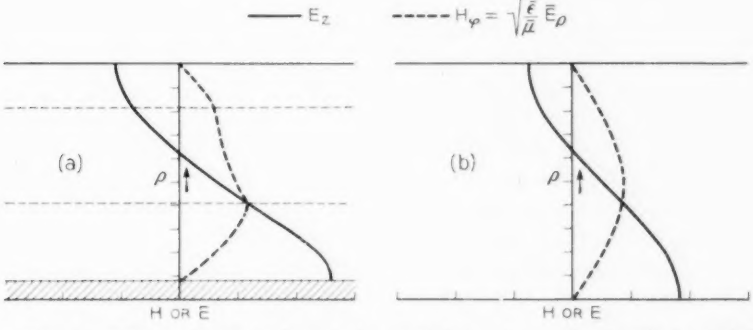


Fig. 14—Fields of principal mode in partially and completely filled coaxial Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

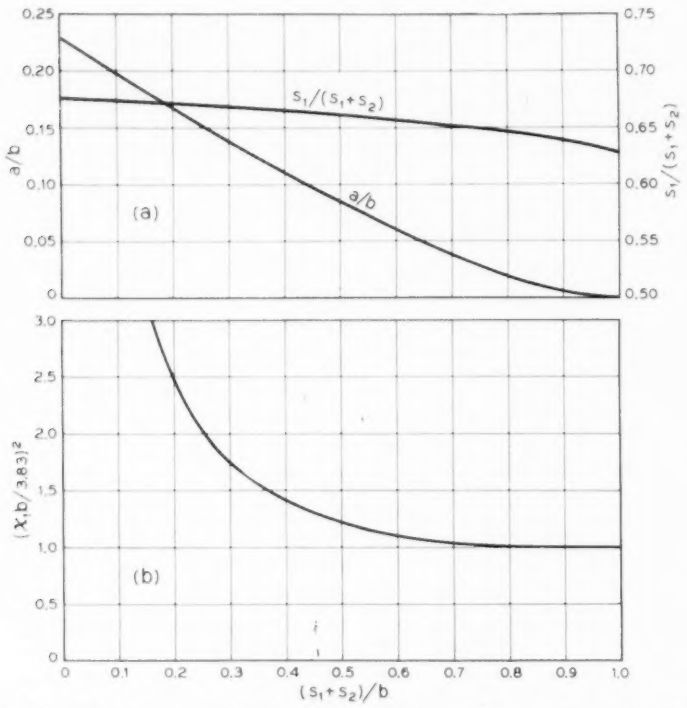


Fig. 15—Relative proportions and relative attenuation constants of optimum Clogston cables with different filling ratios and $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

mum values of these two ratios in the extreme Clogston 1 case, where $(s_1 + s_2) \ll b$, have already been given in equations (138) and (139) of Section IV, while in a complete Clogston 2 with $s_1 + s_2 = b$, the stacks should be divided at the radius $0.6276b$, where according to equation (314) the current density is zero. For intermediate filling ratios, with any fixed magnetic loading ratio $\mu_0/\bar{\mu}$, the optimum distribution of laminated material can most easily be found numerically by calculating χ_1^2 or $(\chi_1 b)^2$ for a number of different choices of the ratios a/b and $s_1/(s_1 + s_2)$, and then locating the minimum by double interpolation.

The results of applying this numerical procedure to Clogston cables with various filling ratios and no magnetic loading are plotted in Fig. 15, the necessary values of $\chi_1 b$ having been found on the analog computer and then refined by desk computation. Fig. 15(a) shows the optimum values of a/b and $s_1/(s_1 + s_2)$ as functions of the filling ratio $(s_1 + s_2)/b$, while Fig. 15(b) shows the corresponding value of $(\chi_1 b/3.83)^2$,

which by equation (365) is proportional to the attenuation constant. We note that the Clogston 2 line with filling ratio unity has the lowest attenuation constant of any cable of the same size without magnetic loading, but that the attenuation constant increases only slowly as the filling ratio decreases, so long as the ratio is greater than about one-half. It also appears that the minimum in $\chi_1 b$, considered as a function of a/b and $s_1/(s_1 + s_2)$ for a fixed filling ratio, is quite broad, which means that in practice one can attain very nearly optimum performance even while deviating somewhat from the optimum proportions.

If the filling ratio is at our disposal, then the solution of the optimum problem is as follows: When there is no magnetic loading of the main dielectric relative to the stacks, that is, when $\mu_0 \leqq \bar{\mu}$, then minimum attenuation is obtained with a complete Clogston 2. If on the other hand there is magnetic loading of the main dielectric, so that $\mu_0 > \bar{\mu}$, then minimum attenuation is obtained with a filling ratio less than unity, whose value is a function of the ratio $\mu_0/\bar{\mu}$.

According to equation (350), the attenuation constant of a plane Clogston line is

$$\alpha = \frac{\chi_1^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}, \tag{374}$$

where χ_1 is given by equation (348),

$$\chi_1 \tan \chi_1 s = \frac{2\bar{\mu}}{\mu_0 b} = \frac{\bar{\mu}}{\mu_0(\frac{1}{2}a - s)}. \tag{375}$$

To find the minimum value of χ_1 when a and $\mu_0/\bar{\mu}$ are given, we differentiate (375) with respect to s and set $d\chi_1/ds$ equal to zero. This gives

$$\chi_1^2 \sec^2 \chi_1 s = \frac{\bar{\mu}}{\mu_0(\frac{1}{2}a - s)^2}, \tag{375.5}$$

which when solved simultaneously with equation (375) leads to

$$\sin \chi_1 s = \sqrt{\bar{\mu}/\mu_0}, \qquad \chi_1 = \frac{1}{s} \sin^{-1} \sqrt{\bar{\mu}/\mu_0}. \tag{376}$$

Substituting this value of χ_1 into (375) and solving for s in terms of a, we get

$$s = \tfrac{1}{2}a \frac{\mu_0 \sin^{-1} \sqrt{\bar{\mu}/\mu_0}}{\mu_0 \sin^{-1} \sqrt{\bar{\mu}/\mu_0} + \sqrt{\bar{\mu}(\mu_0 - \bar{\mu})}}, \tag{377}$$

and from (374) the corresponding attenuation constant is

$$\alpha = \frac{2}{\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2} \left[\sin^{-1} \sqrt{\bar{\mu}/\mu_0} + \sqrt{\bar{\mu}(\mu_0 - \bar{\mu})/\mu_0^2}\right]^2. \tag{378}$$

As $\mu_0/\bar{\mu}$ increases from unity to very large values, the optimum value of s decreases from $\frac{1}{2}a$ toward $\frac{1}{4}a$, so that the filling ratio decreases from unity toward one-half. If $\mu_0/\bar{\mu} < 1$, equation (376) does not yield a real solution, but the complete Clogston 2 is still the physical structure having the lowest attenuation.

For a coaxial Clogston line without magnetic loading the optimum filling ratio is unity, as we have seen above, while in the presence of magnetic loading a smaller filling ratio is optimum. This filling ratio and the optimum distribution of the laminated material in the cable can be determined by numerical analysis for any given value of $\mu_0/\bar{\mu}$. It is reasonably evident on physical grounds, and can be proved mathematically by a variational argument applied to the lowest eigenvalue of equations (369), that whatever may be the radii ρ_1 and ρ_2 of the main dielectric, the lowest attenuation constant is achieved when $a = 0$, that is, when there is no core inside the inner stack. (This is only a mathematical limit; from a practical standpoint, the use of a small core in the manufacturing process is not likely to make any significant increase in the attenuation of the cable.) For each value of the loading ratio $\mu_0/\bar{\mu}$, therefore, we have merely to minimize the value of $(\chi_1 b)^2$ as a function of the two ratios ρ_1/b and ρ_2/b, which can be done by the double interpolation procedure mentioned earlier. We find that as $\mu_0/\bar{\mu}$ increases from unity to very large values, the optimum value of ρ_1 decreases from $0.6276b$ toward $0.3930b$, while ρ_2 increases from $0.6276b$ toward $0.8226b$, so that the filling ratio decreases from unity toward 0.5704. The limiting values of ρ_1 and ρ_2 when $\mu_0/\bar{\mu} \gg 1$ are derived from equation (364) by the method shown in Appendix II.

As a numerical example we have considered a Clogston cable with $\mu_0 = 3\bar{\mu}$. The optimum proportions of this cable and the corresponding value of χ_1 are approximately

$$\rho_1 = 0.426b, \qquad \rho_2 = 0.796b, \qquad \chi_1 = 2.720/b; \tag{379}$$

and the minimum attenuation constant is

$$\alpha = \frac{3.699}{\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}b^2}. \tag{380}$$

The attenuation constant of a complete Clogston 2 with the same stack parameters $\bar{\mu}$ and $\bar{\epsilon}$ is, from (318),

$$\alpha = \frac{7.341}{\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}b^2}, \tag{381}$$

so that the attenuation constant of the optimum loaded cable is only

about 0.504 times that of the optimum unloaded one. In this example, of course, we have said nothing about the effects of magnetic dissipation.

In the above work we have assumed that the electrical constants $\bar{\mu}$, $\bar{\epsilon}$, $\bar{g}$ of the stacks and μ_0 , ϵ_0 of the main dielectric were all fixed quantities. We now consider the case in which the electrical constants of the conducting and insulating layers are given, but the fraction θ of conducting material in the stacks may be varied. We also suppose that the constants of the main dielectric are at our disposal, so that Clogston's condition may always be satisfied. When then is the optimum value of θ?

If we express $\bar{\epsilon}$, $\bar{\mu}$, and $\bar{g}$ in terms of the constants of the individual layers by equations (268) of Section VIII, we find that the expression for the attenuation constant of the principal mode in a Clogston line becomes

$$\alpha = \frac{\chi_1^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}} = \frac{\sqrt{\epsilon_2}\,\chi_1^2}{2\theta(1-\theta)^{\frac{1}{2}}[\theta\mu_1 + (1-\theta)\mu_2]^{\frac{1}{2}} g_1}, \tag{382}$$

where χ_1 is the lowest root of equation (348) for a plane line or equation (364) for a coaxial cable. We wish to minimize α as a function of θ.

If the conducting and insulating layers have different permeabilities ($\mu_1 \neq \mu_2$), then in the general partially filled line χ_1 depends on θ, through the factor $\bar{\mu}$ in equation (348) or (364), as well as on the geometric proportions of the line. In the limiting case of an extreme Clogston 1 line we found in Section IV, equation (145), that the optimum value of θ is

$$\theta = \frac{\mu_1 + (\mu_1^2 + 8\mu_1\mu_2)^{\frac{1}{2}}}{3\mu_1 + (\mu_1^2 + 8\mu_1\mu_2)^{\frac{1}{2}}}; \tag{383}$$

while in a Clogston 2 with no main dielectric, it turns out from (348) or (364) that χ_1 is independent of θ, and an elementary calculation shows that the value of θ which minimizes α is

$$\theta = \frac{3(\mu_1 - 2\mu_2) + (9\mu_1^2 - 4\mu_1\mu_2 + 4\mu_2^2)^{\frac{1}{2}}}{8(\mu_1 - \mu_2)}. \tag{384}$$

For the general partially filled line, however, there seems to be no simple expression for the optimum value of θ.

If the conducting and insulating layers have equal permeabilities, then the average permeability $\bar{\mu}$ ($= \mu_1 = \mu_2$) is independent of θ, and matters are much simpler. Since χ_1 is also independent of θ, the minimum value of α in equation (382) is achieved when

$$\theta = 2/3, \tag{385}$$

that is, when *the thickness of the conducting layers is twice the thickness of the insulating layers*. Thus the result obtained in Section IV for extreme Clogston 1 lines is shown to hold for Clogston lines with an arbitrary degree of filling, provided only that the permeabilities of the conducting and insulating layers are equal.

We emphasize that the preceding results apply only when the layers are infinitesimally thin. If the layers are of finite thickness, then the optimum value of θ will be less than that calculated for infinitesimally thin layers. The case of finite layers will be discussed in Section XI.

X. HIGHER MODES IN CLOGSTON LINES

We shall now investigate certain of the higher modes which are possible in Clogston-type transmission lines. As elsewhere in this paper, we shall restrict ourselves to modes having H_x, E_y, E_z or H_ϕ, E_ρ, E_z field components only, and for simplicity we shall assume stacks of infinitesimally thin laminae backed by high-impedance boundaries; but we shall place no restrictions on the relative thicknesses of the stacks and the main dielectric. We shall suppose, however, that the main dielectric always satisfies Clogston's condition. From physical considerations we anticipate the existence of higher modes of two types:

(1) In a partially filled Clogston line containing a finite thickness of main dielectric, there will be a group of modes very similar to the modes which can propagate between perfect conductors when the frequency is high enough to allow one or more field reversals in the space between the conductors. In a Clogston line these modes will have most of their field energy in the main dielectric, and for lack of a better term may be called "dielectric modes". They will all be cut off at sufficiently low frequencies, and for this reason are not likely to be of much engineering importance. The cutoff frequency of any particular dielectric mode is approximately inversely proportional to the thickness of the main dielectric, so that these modes cannot exist in a completely filled Clogston 2.

(2) There will also be a group of modes which are closely bound up with the laminated stacks, and which correspond to one or more current reversals in the stacks themselves; we shall call these the "stack modes".[21] The stack modes will propagate down to zero frequency on either a partially or a completely filled Clogston line. They will have higher attenuation constants than the principal mode, but occasions may arise in which they are of considerable practical importance. We shall therefore consider these modes in some detail in what follows.

[21] The stack modes in plane lines were discussed by Clogston in Reference 1, Sections IV–VI.

As we have seen in the preceding section, the even and odd modes in a plane Clogston line with infinitesimally thin laminae and high-impedance boundaries correspond respectively to the roots of equations (331) and (332), namely

$$\tanh \tfrac{1}{2}\sqrt{i\omega\bar{\epsilon}/\bar{g}}\;\Gamma_\ell b \tanh \Gamma_\ell s = -\frac{\bar{\mu}}{\mu_0}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}}, \tag{386}$$

$$\coth \tfrac{1}{2}\sqrt{i\omega\bar{\epsilon}/\bar{g}}\;\Gamma_\ell b \tanh \Gamma_\ell s = -\frac{\bar{\mu}}{\mu_0}\sqrt{\frac{i\omega\bar{\epsilon}}{\bar{g}}}. \tag{387}$$

In either case the propagation constant γ is related to Γ_ℓ by

$$\gamma^2 = -\omega^2\bar{\mu}\bar{\epsilon} - (i\omega\bar{\epsilon}/\bar{g})\Gamma_\ell^2, \tag{388}$$

and the field components are given by (333) and (334) for the even modes, or by (335) and (336) for the odd modes.

Our first observation relative to equations (386) and (387) is that the right-hand sides of these equations are extremely small compared to unity. Since the right-hand members are of the order of magnitude of $(\omega\bar{\epsilon}/\bar{g})^{\frac{1}{2}}$, at least one of the two factors on the left side of each equation must be of the order of $(\omega\bar{\epsilon}/\bar{g})^{\frac{1}{4}}$, which is still small compared to unity. If we consider the factors separately, there will be an infinite number of values of Γ_ℓ for which each vanishes, since the hyperbolic tangent vanishes whenever its argument is equal to $m\pi i$, where m is any integer, and the hyperbolic cotangent vanishes whenever its argument equals $(m + \frac{1}{2})\pi i$. Since the coefficients of Γ_ℓ in the two factors on the left side of either equation have different phase angles, we see that both factors cannot vanish simultaneously for any non-zero value of Γ_ℓ. However as we have noted earlier the coefficient of Γ_ℓ in the first factor is very much smaller than the coefficient of Γ_ℓ in the second factor, and so in equation (386) both hyperbolic tangents may be small in the neighborhood of the first few non-zero roots of the second one. On the other hand the second hyperbolic tangent will not be small in the neighborhood of the non-zero roots of the first one; and in equation (387) the hyperbolic tangent and cotangent will never be small simultaneously. With these remarks in mind we shall proceed to a more detailed study of the various higher modes.

One group of modes is given to a good approximation by the condition that the first factor on the left side of equation (386) or (387) vanishes, that is,

$$\sqrt{i\omega\bar{\epsilon}/\bar{g}}\;\Gamma_\ell b \approx m\pi i, \tag{389}$$

where $m = 1, 2, 3, \cdots$, and the even values of m correspond to the even

modes while the odd values correspond to the odd modes. In this section we shall exclude the case $m = 0$, which corresponds to the principal mode discussed in the preceding section. From (389) and equation (330) of Section IX, we get

$$\Gamma_t \approx \frac{m\pi}{b}\sqrt{\frac{i\bar{g}}{\omega\bar{\epsilon}}} = \frac{(1+i)m\pi}{\sqrt{2}b}\sqrt{\frac{\bar{g}}{\omega\bar{\epsilon}}}, \tag{390}$$

$$\kappa_0 \approx \frac{m\pi i}{b}. \tag{391}$$

The fields of the mth mode are given by substituting these quantities into (333) and (334) if m is even, or (335) and (336) if m is odd.

From equation (388), making use of Clogston's condition, the propagation constant of the mth mode of this family is given by

$$\gamma^2 \approx -\omega^2\mu_0\epsilon_0 + m^2\pi^2/b^2 = -4\pi^2/\lambda_0^2 + m^2\pi^2/b^2, \tag{392}$$

where λ_0 is the wavelength of a free wave in the main dielectric at the operating frequency. To this approximation the values of γ are the same as the propagation constants of the family of modes (with H_x, E_y, and E_z field components only) which are possible in a dielectric slab of thickness b between perfectly conducting planes. The cutoff wavelength of the mth mode is

$$\lambda_c = 2b/m, \tag{393}$$

the propagation constant being real, to the present approximation, if $\lambda_0 > \lambda_c$ and pure imaginary if $\lambda_0 < \lambda_c$. We see that the cutoff frequency is inversely proportional to the width of the main dielectric, so that this family of modes is not possible in a completely filled Clogston 2.

It is worth noting that the effective skin depth of the stacks for the mth mode is, from (390),

$$\Delta = \frac{1}{\text{Re }\Gamma_t} = \frac{b}{m\pi}\sqrt{\frac{2\omega\bar{\epsilon}}{\bar{g}}} = \frac{\lambda_c}{\lambda_0}\sqrt{\frac{2}{\omega\bar{\mu}\bar{g}}}. \tag{394}$$

If the mode is just above cutoff, then Δ is of the order of magnitude of δ_1 $(= \sqrt{2/\omega\mu_1 g_1})$, but as ω increases indefinitely Δ also increases indefinitely, for the ideal stack of infinitesimally thin laminae. The physical explanation is simple: When the mode is near cutoff the phase velocity is very high, but as the frequency is increased the phase velocity approaches the velocity of a free wave in the main dielectric, for which the effective skin depth of the stacks was designed by Clogston's condi-

tion to be infinite. By increasing the $\mu_0\epsilon_0$ product of the main dielectric, it would be possible to make the effective skin thickness of a stack of infinitesimally thin layers infinite for any given mode at any single specified frequency, but at the moment this possibility appears of scarcely more than academic interest. Of course the practical limitation on effective skin depth at high frequencies is the finite thickness of the layers, a consideration which we do not take into account in the present section.

The attenuation constants of the dielectric modes, when these modes are above cutoff, may be calculated either by obtaining the small corrections to the values of Γ_ℓ due to the fact that the right side of equation (386) or (387) is not rigorously zero, or by taking one-half the ratio of dissipated power per unit length to transmitted power. Either method gives for the mth mode, assuming the stack thickness s to be large compared to Δ,

$$\alpha = \frac{m\pi}{b^2}\frac{\bar{\mu}}{\mu_0}\frac{\sqrt{2/\omega\bar{\mu}\bar{g}}}{\sqrt{1 - (\lambda_0/\lambda_c)^2}}. \tag{395}$$

Equation (395) assumes conducting layers very thin compared to the skin depth, a situation which may be difficult to achieve at frequencies high enough to permit the modes of this family to propagate.

Another family of modes which can exist on a parallel-plane Clogston line is given by the condition that the second factor on the left side of equation (386) or (387) shall be nearly equal to zero. As pointed out above the even modes present a slight complication; since the coefficient of Γ_ℓ in the first hyperbolic tangent on the left side of (386) is very small, this factor may be comparable to or smaller than the term on the right side in the neighborhood of the first few roots of the equation, in which case the second hyperbolic tangent will not be small compared to unity at these roots. For all the modes in which we can conceivably be interested, however, $|\Gamma_\ell b|$ will be a small fraction of the very large number $2\sqrt{\bar{g}/\omega\bar{\epsilon}}$, and we may therefore replace the first hyperbolic tangent on the left side of (386) by its argument. Thus on making the usual substitution,

$$\Gamma_\ell^2 = -\chi^2, \qquad \Gamma_\ell = i\chi, \tag{396}$$

we get for the even modes,

$$\chi s \tan \chi s = \frac{\bar{\mu}}{\mu_0}\frac{2s}{b}, \tag{397}$$

which is the same as equation (348) of the preceding section. On the

other hand, the odd modes of this family are all given approximately by

$$\tan \chi s = 0, \tag{398}$$

since the hyperbolic cotangent on the left side of (387) will never be small for the same value of Γ_ℓ (or χ) as the hyperbolic tangent.

Equation (397) has an infinite number of positive real roots, which may be denoted by

$$\chi_1 , \chi_3 , \chi_5 , \cdots , \chi_{2p+1} , \cdots , \tag{399}$$

and it is clear that

$$p\pi/s < \chi_{2p+1} \leqq (p + \tfrac{1}{2})\pi/s. \tag{400}$$

If the thickness b of the main dielectric is not zero, so that the right side of equation (397) is finite, the higher roots χ_{2p+1} approach nearer and nearer to $p\pi/s$ as p increases; but if $b = 0$, then

$$\chi_{2p+1} = (p + \tfrac{1}{2})\pi/s = (2p + 1)\pi/a \tag{401}$$

for all p. The positive roots of equation (398) may be called

$$\chi_2 , \chi_4 , \chi_6 , \cdots , \chi_{2p} , \cdots , \tag{402}$$

where

$$\chi_{2p} = p\pi/s \tag{403}$$

for all p; and both sets of roots may be combined in the single sequence

$$\chi_1 , \chi_2 , \chi_3 , \cdots , \chi_p , \cdots . \tag{404}$$

The advantages of designating the principal mode as the first rather than the zero-th mode seem to outweigh the minor disadvantage that in the sequence (404) the odd subscripts correspond to what we have been calling the even modes, and vice versa.

The attenuation and phase constants of the pth mode are obtained in terms of χ_p from (388) and (396). Under the usual assumption that the attenuation per radian is small, we have

$$\alpha = \frac{\chi_p^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}, \tag{405}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}, \tag{406}$$

which become, for the completely filled Clogston 2,

$$\alpha = \frac{p^2\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}, \tag{407}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}}. \tag{408}$$

From (330) and (396) we have for the pth mode,

$$\Gamma_\ell = i\chi_p , \tag{409}$$

$$\kappa_0 = (-1 + i)\sqrt{\omega\tilde{\epsilon}/2\tilde{g}}\ \chi_p . \tag{410}$$

The fields may be obtained by substituting Γ_ℓ and κ_0 into equations (333) and (334) when p is odd, or (335) and (336) when p is even. For the modes in which we are interested, that is, for sufficiently small values of p, we may replace sh $\kappa_0 y$ by $\kappa_0 y$ and ch $\kappa_0 y$ by unity when $|\, y \,| \leqq \frac{1}{2}b$. Then for the modes with odd subscripts $2p + 1$ we have, in the main dielectric,

$$\begin{aligned} H_x &\approx H_0 e^{-\gamma_{2p+1} z}, \\ E_y &\approx -\sqrt{\frac{\bar{\mu}_0}{\epsilon_0}}\, H_0 e^{-\gamma_{2p+1} z}, \\ E_z &\approx \frac{\mu_0 \chi_{2p+1}^2}{\bar{\mu}\tilde{g}}\, H_0 e^{-\gamma_{2p+1} z}, \end{aligned} \tag{411}$$

for $-\frac{1}{2}b \leqq y \leqq \frac{1}{2}b$, while in the stacks,

$$\begin{aligned} H_x &\approx H_0 \frac{\sin \chi_{2p+1}(\frac{1}{2}a \mp y)}{\sin \chi_{2p+1} s}\, e^{-\gamma_{2p+1} z}, \\ \tilde{E}_y &\approx -\sqrt{\frac{\bar{\mu}}{\tilde{\epsilon}}}\, H_0 \frac{\sin \chi_{2p+1}(\frac{1}{2}a \mp y)}{\sin \chi_{2p+1} s}\, e^{-\gamma_{2p+1} z}, \\ E_z &\approx \pm \frac{\chi_{2p+1}}{\tilde{g}}\, H_0 \frac{\cos \chi_{2p+1}(\frac{1}{2}a \mp y)}{\sin \chi_{2p+1} s}\, e^{-\gamma_{2p+1} z}, \end{aligned} \tag{412}$$

for $\frac{1}{2}b \leqq |\, y \,| \leqq \frac{1}{2}a$, where the upper signs refer to the upper stack and the lower signs to the lower stack. Of course the arbitrary amplitude factor H_0 need not be the same for different values of p. Similarly for the modes with even subscripts $2p$ the fields in the main dielectric are

$$\begin{aligned} H_x &\approx 0, \\ E_y &\approx 0, \\ E_z &\approx (-)^p \frac{p\pi}{\tilde{g}s}\, H_0 e^{-\gamma_{2p} z}, \end{aligned} \tag{413}$$

for $-\frac{1}{2}b \leqq y \leqq \frac{1}{2}b$, while in the stacks,

$$H_x \approx \pm H_0 \sin \frac{p\pi(\frac{1}{2}a \mp y)}{s} e^{-\gamma_{2p}z},$$

$$\bar{E}_y \approx \mp \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}} H_0 \sin \frac{p\pi(\frac{1}{2}a \mp y)}{s} e^{-\gamma_{2p}z}, \tag{414}$$

$$E_z \approx \frac{p\pi}{\bar{g}s} H_0 \cos \frac{p\pi(\frac{1}{2}a \mp y)}{s} e^{-\gamma_{2p}z},$$

for $\frac{1}{2}b \leqq |y| \leqq \frac{1}{2}a$, and again the upper signs refer to the upper stack and the lower signs to the lower stack.

In a complete Clogston 2 the expressions for the fields simplify a good deal. For the modes with odd subscripts $2p + 1$ the fields are, for $-\frac{1}{2}a \leqq y \leqq \frac{1}{2}a$,

$$H_x \approx H_0 \cos \frac{(2p+1)\pi y}{a} e^{-\gamma_{2p+1}z},$$

$$\bar{E}_y \approx -\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}} H_0 \cos \frac{(2p+1)\pi y}{a} e^{-\gamma_{2p+1}z}, \tag{415}$$

$$E_z \approx \frac{(2p+1)\pi}{\bar{g}a} H_0 \sin \frac{(2p+1)\pi y}{a} e^{-\gamma_{2p+1}z},$$

while for the modes with even subscripts $2p$,

$$H_x \approx H_0 \sin \frac{2p\pi y}{a} e^{-\gamma_{2p}z},$$

$$\bar{E}_y \approx -\sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}} H_0 \sin \frac{2p\pi y}{a} e^{-\gamma_{2p}z}, \tag{416}$$

$$E_z \approx -\frac{2p\pi}{\bar{g}a} H_0 \cos \frac{2p\pi y}{a} e^{-\gamma_{2p}z}.$$

The fields of the higher modes in a plane Clogston 2 are simply related to the fields of the principal mode shown in Fig. 13(d). The fields of the pth mode may be obtained conceptually by stacking up p "layers", each of thickness a/p, the fields in each layer being identical with the fields of the principal mode except for the scale reduction and a phase difference of 180° between adjacent layers. Equation (407) shows that the attenuation constant of the pth mode in a plane Clogston 2 with infinitesimally thin laminae and high-impedance walls is just p^2 times the attenuation constant of the principal mode.

It may be observed that if we are considering a partially filled plane Clogston line with $b > 0$, then the propagation constants of the $2p$th

and the $(2p + 1)$st stack modes will be nearly the same for sufficiently large values of p (how large depends on the ratio of stack thickness to main dielectric thickness). Except for differences in sign, the fields in the stacks will also be the same up to second order differences which our approximations do not show. The physical interpretation is that for a thick enough main dielectric and/or sufficiently large values of p, the fields are confined almost entirely to the two stacks, being relatively small in the main dielectric, while the stacks act like a pair of almost independent Clogston 2 lines each of thickness s and carrying a particular Clogston 2 higher mode.

Figs. 16 and 17 show field plots for the second and third stack modes (i.e., the first and second higher modes) in the same four plane Clogston lines that were used to exhibit the behavior of the principal mode in Fig. 13. Note however that these plots are not normalized and that the horizontal scales on the figures are not all the same. The following table gives $\chi_2 s$ and $\chi_3 s$ as functions of the fraction $s/\frac{1}{2}a$ of the total space filled by the stacks, and also the quantities $(\chi_2 a/\pi)^2$ and $(\chi_3 a/\pi)^2$, which are just the ratios of the attenuation constants of

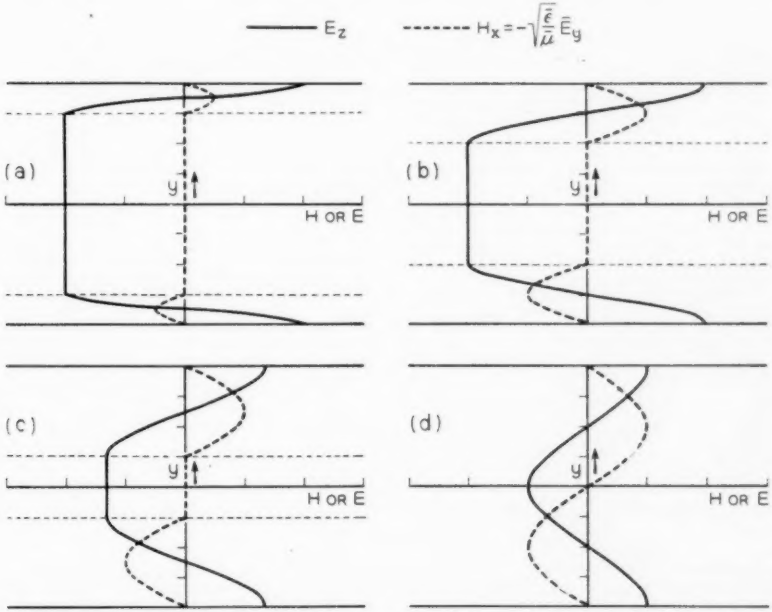


Fig. 16—Fields of second stack mode in partially and completely filled plane Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

these modes to the attenuation constant of the principal mode in a completely filled Clogston 2.

$\rho_1/\frac{1}{2}\rho_2$	χ_{21}	χ_{31}	$(\chi_{21}/\pi)^2$	$(\chi_{31}/\pi)^2$
$\frac{1}{4}$	π	3.244	64.0	68.2
$\frac{1}{2}$	π	3.426	16.0	19.0
$\frac{3}{4}$	π	3.809	7.1	10.5
1	π	4.712	4.0	9.0

These results may be compared with those given in Section IX for the principal mode in plane Clogston lines having the same proportions.

Turning now to the cylindrical geometry, we remember that all the circular transverse magnetic modes in a coaxial Clogston line with infinitesimally thin laminae and high-impedance boundaries are given by the roots of equation (339) of Section IX, namely

$$\frac{\kappa_0 K_0(\kappa_0\rho_1) + i\omega\epsilon_0 Z_1 K_1(\kappa_0\rho_1)}{\kappa_0 I_0(\kappa_0\rho_1) - i\omega\epsilon_0 Z_1 I_1(\kappa_0\rho_1)} = \frac{\kappa_0 K_0(\kappa_0\rho_2) - i\omega\epsilon_0 Z_2 K_1(\kappa_0\rho_2)}{\kappa_0 I_0(\kappa_0\rho_2) + i\omega\epsilon_0 Z_2 I_1(\kappa_0\rho_2)}, \quad (417)$$

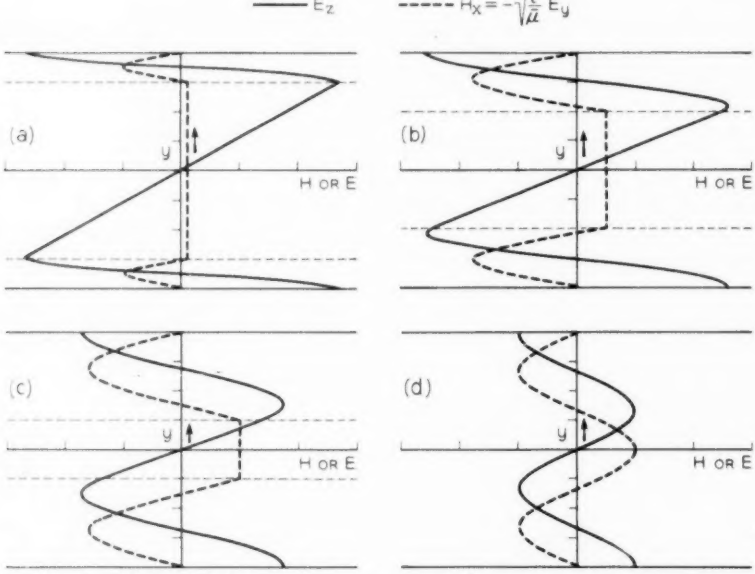


Fig. 17—Fields of third stack mode in partially and completely filled plane Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

where

$$Z_1 = \frac{\Gamma_\ell}{\tilde{g}} \frac{K_0(\Gamma_\ell \rho_1) I_1(\Gamma_\ell a) + K_1(\Gamma_\ell a) I_0(\Gamma_\ell \rho_1)}{K_1(\Gamma_\ell a) I_1(\Gamma_\ell \rho_1) - K_1(\Gamma_\ell \rho_1) I_1(\Gamma_\ell a)}, \tag{418}$$

$$Z_2 = \frac{\Gamma_\ell}{\tilde{g}} \frac{K_0(\Gamma_\ell \rho_2) I_1(\Gamma_\ell b) + K_1(\Gamma_\ell b) I_0(\Gamma_\ell \rho_2)}{K_1(\Gamma_\ell \rho_2) I_1(\Gamma_\ell b) - K_1(\Gamma_\ell b) I_1(\Gamma_\ell \rho_2)}. \tag{419}$$

Physically it is clear that the modes of the coaxial cable must be of the same general types as the modes of the parallel-plane line, and so in seeking the roots of equation (417) we shall be guided by the results which we have already found for the plane structure.

The dielectric modes in the cable may be located, to a first approximation, by setting Z_1 and Z_2 equal to zero, whence (417) becomes

$$\frac{K_0(\kappa_0 \rho_1)}{I_0(\kappa_0 \rho_1)} = \frac{K_0(\kappa_0 \rho_2)}{I_0(\kappa_0 \rho_2)}. \tag{420}$$

The substitution

$$\kappa_0^2 = -h^2, \qquad \kappa_0 = ih, \tag{421}$$

transforms (420) into

$$J_0(h\rho_1) N_0(h\rho_2) - J_0(h\rho_2) N_0(h\rho_1) = 0. \tag{422}$$

Equation (422) has an infinite number of real roots h_1, h_2, h_3, $\cdots$, h_m, $\cdots$, of which the mth one may be written in the form[22]

$$h_m = \frac{m\pi F_m(\rho_1/\rho_2)}{\rho_2 - \rho_1}, \tag{423}$$

where $F_m(\rho_1/\rho_2)$ is a function which increases from slightly less than unity at $\rho_1/\rho_2 = 0$ to unity at $\rho_1/\rho_2 = 1$. From equations (330) and (423) we have, approximately,

$$\Gamma_\ell = h_m \sqrt{\frac{i\tilde{g}}{\omega\tilde{\epsilon}}} = \frac{(1+i) m\pi F_m(\rho_1/\rho_2)}{\sqrt{2}\,(\rho_2 - \rho_1)} \sqrt{\frac{\tilde{g}}{\omega\tilde{\epsilon}}}, \tag{424}$$

$$\kappa_0 = ih_m = \frac{im\pi F_m(\rho_1/\rho_2)}{\rho_2 - \rho_1}, \tag{425}$$

and the fields of the mth mode are given by substituting these expressions into equations (340) to (344) of the preceding section.

From equation (388) the propagation constant of the mth dielectric

[22] Reference 18, pp. 204–206. What we call $m\pi F_m(\rho_1/\rho_2)$ is tabulated by Jahnke and Emde, pp. 205–206, as $(k-1)x_0^{(m)}$, where $k = \rho_2/\rho_1$.

mode is defined by

$$\gamma^2 = -\omega^2\mu_0\epsilon_0 + h_m^2 = -4\pi^2/\lambda_0^2 + h_m^2 \tag{426}$$

to the present approximation, and the cutoff wavelength is

$$\lambda_c = \frac{2\pi}{h_m} = \frac{2(\rho_2 - \rho_1)}{mF_m(\rho_1/\rho_2)}, \tag{427}$$

which tends to zero with the thickness $\rho_2 - \rho_1$ of the main dielectric.

As in the parallel-plane case, when the mth dielectric mode is just able to propagate the effective skin depth in the stacks is of the order of δ_1, and the stack impedances are approximately

$$Z_1 = Z_2 = K = \Gamma_\ell/\bar{g}, \tag{428}$$

under the present assumption of infinitesimally thin laminae. The power dissipated in the stacks and the corresponding attenuation constant may be calculated by a straightforward procedure if desired.

Before leaving the subject of higher dielectric modes in a Clogston cable, we should point out that although we have mentioned only the transverse magnetic modes with circular symmetry, in reality there exist a double infinity of both transverse magnetic and transverse electric higher modes. These modes are discussed in textbooks[23] for coaxial lines bounded by perfect conductors, and they will propagate, with minor changes due to wall losses, in either ordinary or Clogston-type coaxial cables if the frequency is high enough. At ordinary engineering frequencies, however, the higher modes contribute only to the local fields excited at discontinuities, and are therefore not of any great practical importance.

To find the stack modes in a Clogston cable we assume, subject to a posteriori verification, that in the main dielectric we shall have $|\kappa_0\rho| \ll 1$ for all the modes of interest. Then if we set $\Gamma_\ell = i\chi$, equation (417) reduces, as in Section IX, to

$$\frac{1}{\chi\rho_1}\frac{J_1(\chi a)N_0(\chi\rho_1) - N_1(\chi a)J_0(\chi\rho_1)}{J_1(\chi a)N_1(\chi\rho_1) - N_1(\chi a)J_1(\chi\rho_1)} + \frac{1}{\chi\rho_2}\frac{J_1(\chi b)N_0(\chi\rho_2) - N_1(\chi b)J_0(\chi\rho_2)}{J_1(\chi\rho_2)N_1(\chi b) - N_1(\chi\rho_2)J_1(\chi b)} = \frac{\mu_0}{\bar{\mu}}\log\frac{\rho_2}{\rho_1}, \tag{429}$$

which is the same as equation (364). Equation (429) has an infinite

[23] A good account is given by N. Marcuvitz, *Waveguide Handbook*, M. I. T. Rad. Lab. Series, **10**, McGraw-Hill, New York, 1951, pp. 72–80.

number of real roots,

$$\chi_1 , \chi_2 , \chi_3 , \cdots , \chi_p , \cdots , \tag{430}$$

of which χ_1 corresponds to the principal mode and $\chi_2 , \chi_3 , \cdots ,$ to the higher stack modes. The χ's are the eigenvalues of the system of equations (369), and as such may be located approximately with a differential analyzer, or as accurately as desired by numerical solution of equation (429). The attenuation and phase constants of the pth mode are

$$\alpha = \frac{\chi_p^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}} , \tag{431}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}} , \tag{432}$$

provided that the attenuation per radian is small, i.e., that p is not too large. The fields are given by writing χ_p for χ_1 and γ_p for γ in equations (371) to (373) of Section IX.

For a Clogston 2 with no main dielectric we can set $\rho_1 = \rho_2$ in equation (429) and obtain the much simpler form

$$J_1(\chi a)N_1(\chi b) - J_1(\chi b)N_1(\chi a) = 0. \tag{433}$$

The pth root of (433) may be written[24]

$$\chi_p = \frac{p\pi f_p(a/b)}{b - a} , \tag{434}$$

where the functions $f_p(a/b)$ have values slightly greater than unity when $a/b = 0$, and decrease monotonically toward 1 as a/b approaches unity. The attenuation and phase constants of the pth mode in a Clogston 2 are given by

$$\alpha = \frac{p^2\pi^2 f_p^2(a/b)}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}(b - a)^2} , \tag{435}$$

$$\beta = \omega\sqrt{\bar{\mu}\bar{\epsilon}} , \tag{436}$$

provided that p is not too large. The attenuation constant of the pth mode is thus approximately p^2 times the attenuation constant of the principal mode, the approximation being better the closer the ratio a/b is to unity. The fields of the pth mode may be obtained by writing χ_p for χ_1 and γ_p for γ in equations (302) of Section VIII, or equations (311) if $a = 0$. Qualitatively these fields are very similar to the fields of

[24] Reference 18, pp. 204–206. What we call $p\pi f_p(a/b)$ is tabulated by Jahnke and Emde as $(k - 1)x_1^{(p)}$, where $k = b/a$.

the pth mode in a parallel-plane Clogston 2, with the same number of field maxima and field reversals for a given value of p, though of course the spacings and amplitudes of the field maxima are not all equal in the coaxial cable.

As numerical examples we have plotted in Figs. 18 and 19 the fields of the second and third stack modes (i.e., the first and second higher modes) in the same two Clogston cables which were used to show the principal mode in Fig. 14. The horizontal scales on these figures are arbitrary and have no relation to one another. Figs. 18(a) and 19(a) represent a partially filled cable with the same dimensions, namely $a = 0.084b$, $\rho_1 = 0.415b$, and $\rho_2 = 0.831b$, as in Fig. 14(a), while Figs. 18(b) and 19(b) represent a completely filled cable, as in Fig. 14(b). The following table shows, as a function of the filling ratio $(s_1 + s_2)/b$, the quantity $(\chi_p b/3.83)^2$ for $p = 1, 2, 3$; this quantity is just the ratio of the attenuation constant of the given mode to the attenuation constant of the principal mode in a completely filled Clogston 2.

$(s_1 + s_2)/b$	$(\chi_1 b/3.83)^2$	$(\chi_2 b/3.83)^2$	$(\chi_3 b/3.83)^2$
0.5	1.234	8.369	24.273
1.0	1.000	3.352	7.050

We note that although the proportions of the partially filled cable were found in Section IX to be optimum, in the sense of minimizing the attenuation constant, for the principal mode in a cable with filling ratio 0.5, there is no reason to believe that the same proportions will be optimum for the second and third modes with the same filling ratio.

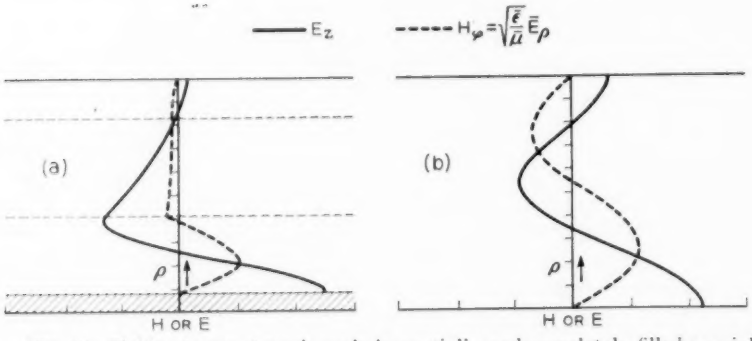


Fig. 18—Fields of second stack mode in partially and completely filled coaxial Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

XI. EFFECT OF FINITE LAMINA THICKNESS. FREQUENCY DEPENDENCE OF ATTENUATION IN CLOGSTON 2 LINES

We shall now study Clogston 2 lines with laminae of finite thickness, and shall investigate the important practical question of how the propagation constant varies with frequency in such lines. Much of the analysis of the present section will deal with parallel-plane structures, but we may be confident that the results will also give at least a good qualitative estimate of the behavior of coaxial cables.

The notation for the plane Clogston 2, shown in Fig. 10, is the same as before, except that we now assume the thicknesses of the individual conducting and insulating layers to be t_1 and t_2 respectively. For definiteness we shall suppose that there are $2n$ conducting layers and $2n$ insulating layers in the whole stack, with a conducting layer next to the lower sheath and an insulating layer next to the upper sheath, though the precise arrangement is of no real importance if the number of layers is large. The total thickness a of the stack is $2n(t_1 + t_2)$, and the fraction of conducting material will as usual be called θ.

The boundary conditions for any mode (having H_x, E_y, and E_z field components only) require that the sum of the impedances looking in opposite directions normal to any plane y = constant be zero. If we match impedances at the lower sheath $y = -\frac{1}{2}a$ and use equation (65) of Section III for the impedance looking into the stack, we have

$$\frac{\frac{1}{2}Z_n(\gamma)(K_1e^{2n\Gamma} + K_2e^{-2n\Gamma}) + K_1K_2 \operatorname{sh} 2n\Gamma}{Z_n(\gamma) \operatorname{sh} 2n\Gamma + \frac{1}{2}(K_1e^{-2n\Gamma} + K_2e^{2n\Gamma})} + Z_n(\gamma) = 0, \quad (437)$$

where Γ, K_1, and K_2 are given by equations (61) and (63). If equation

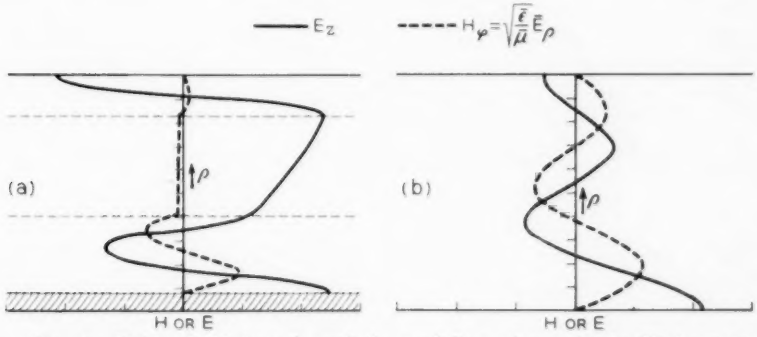


Fig. 19—Fields of third stack mode in partially and completely filled coaxial Clogston lines with $\mu_0 = \bar{\mu}$, $\epsilon_0 = \bar{\epsilon}$.

(437) is solved for $e^{4n\Gamma}$, it takes the form

$$e^{4n\Gamma} = \frac{[Z_n(\gamma) - K_1][Z_n(\gamma) - K_2]}{[Z_n(\gamma) + K_1][Z_n(\gamma) + K_2]}. \tag{438}$$

In order to simplify the general expressions for Γ, K_1, and K_2, we make the following approximations:

(i) We neglect $\omega\epsilon/g_1$ compared to unity, where ϵ represents the dielectric constant of either the conducting or the insulating layers. As we have said before, this is an exceedingly good approximation at all engineering frequencies.

(ii) We neglect γ^2/σ_1^2 $(= \gamma^2/i\omega\mu_1 g_1)$ compared to unity. It turns out that not only is this approximation valid in the frequency range of greatest interest, where γ is approximately equal to $i\omega\sqrt{\bar{\mu}\bar{\epsilon}}$, but also it is valid all the way down to zero frequency, so that in the present section we can easily derive results for the complete frequency range down to dc. So long as $\gamma^2/\sigma_1^2 \ll 1$, we have from equations (56) of Section III,

$$\kappa_1 \approx \sigma_1 , \qquad \eta_{1y} \approx \eta_1 . \tag{439}$$

(iii) We suppose that the thickness t_2 of the layers of insulation is so small that we may replace sh $\kappa_2 t_2$ by $\kappa_2 t_2$ and ch $\kappa_2 t_2$ by unity. These approximations will be amply justified if t_2 is not greater than a few times the skin depth in the conducting layers at the highest operating frequency.

The foregoing approximations lead to results identical with equations (86) and (87) of Section III, namely

$$\text{ch } \Gamma = \frac{\eta_{2y}\kappa_2 t_2}{2\eta_{1y}} \text{sh } \kappa_1 t_1 + \text{ch } \kappa_1 t_1 , \tag{440}$$

and

$$\begin{aligned} K_1 &= -\tfrac{1}{2}\eta_{2y}\kappa_2 t_2 + \sqrt{(\tfrac{1}{2}\eta_{2y}\kappa_2 t_2)^2 + \eta_{1y}\eta_{2y}\kappa_2 t_2 \coth \kappa_1 t_1 + \eta_{1y}^2} , \\ K_2 &= +\tfrac{1}{2}\eta_{2y}\kappa_2 t_2 + \sqrt{(\tfrac{1}{2}\eta_{2y}\kappa_2 t_2)^2 + \eta_{1y}\eta_{2y}\kappa_2 t_2 \coth \kappa_1 t_1 + \eta_{1y}^2} . \end{aligned} \tag{441}$$

To simplify the notation, we introduce the symbol q defined by

$$\begin{aligned} q &= -\frac{\eta_{2y}\kappa_2 t_2}{\eta_1 \sigma_1 t_1} \\ &= -\frac{(1-\theta)\mu_2}{\theta\mu_1}\left[1 + \frac{\gamma^2}{\omega^2\mu_2\epsilon_2}\right] \\ &= 1 - \frac{\bar{\mu}}{\theta\mu_1}\left[1 + \frac{\gamma^2}{\omega^2\bar{\mu}\bar{\epsilon}}\right], \end{aligned} \tag{442}$$

where $\hat{\epsilon}$ and $\hat{\mu}$ are given by equations (268) of Section VIII. The propagation constant γ is thus related to q by

$$\begin{aligned}\gamma &= i\omega\sqrt{\mu_2\epsilon_2}\left[1 + \frac{\theta\mu_1 q}{(1-\theta)\mu_2}\right]^{\frac{1}{2}} \\ &= i\omega\sqrt{\hat{\mu}\hat{\epsilon}}\left[1 + \frac{\theta\mu_1(q-1)}{\hat{\mu}}\right]^{\frac{1}{2}}.\end{aligned} \tag{443}$$

In terms of q and the electrical thickness parameter Θ used in Part I, namely

$$\Theta = \sigma_1 t_1 = (1+i)t_1/\delta_1 \approx \kappa_1 t_1 , \tag{444}$$

equations (440) and (441) become, approximately,

$$\operatorname{ch}\Gamma = \operatorname{ch}\Theta - \tfrac{1}{2}q\Theta \operatorname{sh}\Theta, \tag{445}$$

and

$$\begin{aligned}K_1 &= \frac{\Theta}{g_1 t_1}\left[\tfrac{1}{2}q\Theta + \sqrt{\tfrac{1}{4}q^2\Theta^2 - q\Theta\coth\Theta + 1}\,\right], \\ K_2 &= \frac{\Theta}{g_1 t_1}\left[-\tfrac{1}{2}q\Theta + \sqrt{\tfrac{1}{4}q^2\Theta^2 - q\Theta\coth\Theta + 1}\,\right].\end{aligned} \tag{446}$$

In the general case when the sheath impedance $Z_n(\gamma)$ is a given function of γ, we substitute the expressions for K_1 and K_2 into equation (438), namely

$$e^{4n\Gamma} = \frac{Z_n^2(\gamma) - (K_1+K_2)Z_n(\gamma) + K_1K_2}{Z_n^2(\gamma) + (K_1+K_2)Z_n(\gamma) + K_1K_2}, \tag{447}$$

and then determine γ for each mode by simultaneous numerical solution of equations (443), (445), and (447). At least as a first approximation we may neglect the total current in either sheath compared to the one-way current in the stack; to this approximation $Z_n(\gamma)$ is effectively infinite and (447) becomes

$$e^{4n\Gamma} = 1. \tag{448}$$

The non-zero roots of this equation are

$$\Gamma = ip\pi/2n, \qquad p = 1, 2, 3, \cdots, \tag{449}$$

where $p = 1$ corresponds to the principal mode and the higher values of p to the higher modes discussed in Section X. (We would get nothing new by including negative values of p.) The quantities q and γ for each mode are then given by equations (445) and (443) respectively. If we

wish to take the finite value of $Z_n(\gamma)$ into account, we may calculate K_1 and K_2 from (446) and then obtain a second approximation to Γ from (447); and this process may be repeated as often as desired. From the experience gained in treating a particular example we feel that the method of successive approximations is entirely feasible, but it does involve a considerable amount of numerical work.

In the calculation just described we have to choose the correct sign of the complex square root occurring in the expressions for K_1 and K_2. Without attempting to give a complete discussion of this point here, we observe that it may be shown that

$$\operatorname{sh} \Gamma = \operatorname{sh} \Theta \sqrt{\tfrac{1}{4} q^2 \sigma^2 - q(\Theta \coth \Theta + 1}. \tag{450}$$

Under ordinary circumstances Γ will be a small complex number with a phase angle of about 90°, and Θ will be a small complex number with a phase angle of 45°. Hence the phase angle of the square root may be expected to be in the neighborhood of 45° rather than 225°.

In the remainder of this section we shall restrict ourselves to the case of high-impedance sheaths, so that the values of Γ are given to a sufficiently good approximation by equation (449). We shall discuss the principal mode and the higher modes concurrently, but shall assume throughout that the mode number p is small compared to n. From (445) and (449), the value of q for the pth mode is

$$q = \frac{2\left(\operatorname{ch} \Theta - \cos \frac{p\pi}{2n}\right)}{\Theta \operatorname{sh} \Theta}, \tag{451}$$

and the propagation constant γ is obtained by substituting this value of q into equation (443).

We shall now discuss the variation of the propagation constant of a plane Clogston 2 line with frequency over the full frequency range from zero to very high frequencies.[25] To do this we shall derive approximate expressions for the propagation constant at what may be called, roughly, very low frequencies, low frequencies, high frequencies, and very high frequencies. It will appear presently that the limits of these various frequency ranges depend among other things on the dimensions of the laminated transmission line and the thicknesses of the individual layers, and that the frequency range of greatest engineering importance is what we have called simply "low frequencies".

From equation (444) we have

[25] In this connection see also Reference 1, Appendices A and B, pp. 525–529.

$$\Theta = (1 + i)t_1\sqrt{\omega\mu_1 g_1/2} = (1 + i)t_1\sqrt{\pi\mu_1 g_1 f}, \tag{452}$$

which is proportional to the square root of frequency. For small Θ equation (451) may be written

$$\begin{aligned} q &= \left[2(\text{ch } \Theta - 1) + 4 \sin^2 \frac{p\pi}{4n}\right] \frac{\text{csch } \Theta}{\Theta} \\ &= \left[4 \sin^2 \frac{p\pi}{4n}\right] \frac{1}{\Theta^2} + \left[1 - \frac{2}{3} \sin^2 \frac{p\pi}{4n}\right] \\ &\qquad - \frac{1}{12}\left[1 - \frac{14}{15} \sin^2 \frac{p\pi}{4n}\right]\Theta^2 + \cdots, \end{aligned} \tag{453}$$

on expanding the right side in powers of Θ by Dwight 657.2 and 657.8. If we replace $\sin p\pi/4n$ by $p\pi/4n$ and neglect the square of this quantity in comparison with unity, (453) becomes

$$q \approx \frac{p^2\pi^2}{4n^2} \frac{1}{\Theta^2} + 1 - \frac{\Theta^2}{12} + \cdots. \tag{454}$$

Introducing this expression for q into equation (443) and substituting for Θ from (452), we get for the propagation constant,

$$\gamma = i\omega\sqrt{\bar{\mu}\bar{\epsilon}}\left[1 - \frac{i\theta\mu_1}{\bar{\mu}}\left(\frac{p^2\pi^2}{4n^2\omega\mu_1 g_1 t_1^2} + \frac{\omega\mu_1 g_1 t_1^2}{12}\right)\right]^{\frac{1}{2}}. \tag{455}$$

As the frequency approaches zero in a Clogston 2 line of finite thickness, the term in $1/\omega$ dominates the other terms in square brackets in equation (455). Thus at very low frequencies the attenuation and phase constants of the pth mode are given approximately by

$$\alpha = \frac{p\pi\theta}{2T_1}\sqrt{\frac{\omega\bar{\epsilon}}{2\bar{g}}} = \frac{p\pi}{a}\sqrt{\frac{\pi\bar{\epsilon}f}{\bar{g}}}, \tag{456}$$

$$\beta = \frac{p\pi\theta}{2T_1}\sqrt{\frac{\omega\bar{\epsilon}}{2\bar{g}}} = \frac{p\pi}{a}\sqrt{\frac{\pi\bar{\epsilon}f}{\bar{g}}}, \tag{457}$$

where $2T_1$ $(= 2nt_1)$ is the total thickness of conducting material in the stack of thickness a. To this approximation the attenuation and phase constants are equal, and are proportional to the square root of frequency We note that at very low frequencies,

$$\frac{\gamma^2}{\sigma_1^2} = \frac{2ip^2\pi^2\theta^2\omega\bar{\epsilon}}{8T_1^2\bar{g}} \cdot \frac{1}{i\omega\mu_1 g_1} = \frac{\theta p^2\pi^2\bar{\epsilon}/\mu_1}{a^2\bar{g}^2}, \tag{458}$$

which will be very small compared to unity for lines of all reasonable dimensions, in agreement with our assumption (ii) above.

As the frequency is increased there will be a range in which the terms in parentheses in equation (455) are small compared to unity, so that the square root may be expanded by the binomial theorem. This gives

$$\alpha = \frac{p^2\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2} + \frac{\omega^2\mu_1^2\bar{g}t_1^2}{24\sqrt{\bar{\mu}/\bar{\epsilon}}}, \tag{459}$$

$$\beta = i\omega\sqrt{\bar{\mu}\bar{\epsilon}}\,. \tag{460}$$

If the line is of finite total thickness a and the frequency is so low or the laminae are so thin that the first term on the right side of (459) is large compared to the second, we have approximately

$$\alpha = \frac{p^2\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}. \tag{461}$$

This is the frequency-independent attenuation constant that we found in Section X, equation (407), for the pth mode in a plane Clogston 2 with infinitesimally thin laminae. We shall call the range over which the attenuation is essentially flat the "low-frequency" range. On the other hand, if the laminae are of finite thickness the second term on the right side of (459) ultimately becomes dominant, and the attenuation constant is then given approximately by

$$\alpha = \frac{\omega^2\mu_1^2\bar{g}t_1^2}{24\sqrt{\bar{\mu}/\bar{\epsilon}}} = \frac{\pi^2\mu_1^2\bar{g}t_1^2f^2}{6\sqrt{\bar{\mu}/\bar{\epsilon}}}. \tag{462}$$

This is also the attenuation constant of a plane wave in an unbounded laminated medium (except at very high frequencies), as may be seen by letting the stack thickness a tend to infinity in equation (459). By "high frequencies" we shall mean the frequency range in which the attenuation constant is approximately proportional to f^2.

Finally at very high frequencies when $|\Theta| \gg 1$, we have from (451),

$$q \approx 2/\Theta\,, \tag{463}$$

and so from (443),

$$\gamma = i\omega\sqrt{\mu_2\epsilon_2}\left[1 + \frac{2\theta\mu_1}{(1-\theta)\mu_2\Theta}\right]^{\frac{1}{2}}. \tag{464}$$

Expanding by the binomial theorem and substituting for Θ from (452), we get after a little rearrangement,

$$\alpha = \frac{1}{\sqrt{\mu_2/\epsilon_2}\, t_2 g_1 \delta_1} = \frac{1}{\sqrt{\mu_2/\epsilon_2}\, t_2} \sqrt{\frac{\pi \mu_1 f}{g_1}}, \tag{465}$$

$$\beta = \omega \sqrt{\mu_2 \epsilon_2} + \frac{1}{\sqrt{\mu_2/\epsilon_2}\, t_2 g_1 \delta_1}. \tag{466}$$

Comparing these expressions with equations (25) and (26) of Section II, we see that they correspond to waves in parallel-plane transmission lines of width t_2, bounded by electrically thick solid conductors. We shall call this range, in which α is proportional to the square root of frequency, the "very high frequency" range. At these frequencies the propagation constant is the same in a Clogston 2 line of finite thickness as in an infinite laminated medium.

In order to describe the various frequency ranges more precisely, we shall define the three critical frequencies f_1', f_2', and f_3' to be the frequencies at which the approximate expressions for the attenuation constants in two adjacent frequency ranges are equal. These frequencies are closely related to the critical frequencies which we defined in equations (178) of Section V, when we were discussing the surface impedance of a plane stack of finite layers. For a stack containing a total thickness T_1 of conducting material, we recall that the critical frequencies were f_1, where $T_1 = \delta_1$; f_2, where $T_1 = T_\Delta$; and f_3, where $t_1 = \sqrt{3}\,\delta_1$. For the pth mode in a Clogston 2 containing a total thickness $2T_1$ of conducting material, the frequencies turn out to be

$$\begin{aligned} f_1' &= \frac{p^2 \pi \theta}{16 \bar{\mu} g_1 T_1^2} = \frac{p^2 \pi^2 \theta \mu_1}{16 \bar{\mu}} f_1, \\ f_2' &= \frac{\sqrt{3}\, p}{2 \mu_1 g_1 t_1 T_1} = \frac{p\pi}{2} f_2, \\ f_3' &= \left[\frac{36 \bar{\mu}}{(1-\theta)\mu_2}\right]^{\frac{1}{2}} \frac{1}{\pi \mu_1 g_1 t_1^2} = \left[\frac{4 \bar{\mu} \bar{\epsilon}}{3 \mu_2 \epsilon_2}\right]^{\frac{1}{2}} f_3, \end{aligned} \tag{467}$$

where of course $p = 1$ for the principal mode. Thus the attenuation constant is given approximately by (456) for $0 \leqq f \leqq f_1'$, by (461) for $f_1' \leqq f \leqq f_2'$, by (462) for $f_2' \leqq f \leqq f_3'$, and by (465) for $f \geqq f_3'$.

If we plot the foregoing approximate expressions for the attenuation constant against frequency on log-log paper, we can get a good idea of the variation of the attenuation of a Clogston 2 over the entire frequency range. Both the approximate expressions and the exact results for a particular numerical case are plotted in Fig. 20, for a Clogston 2 of finite thickness and also for an infinite laminated medium. The actual

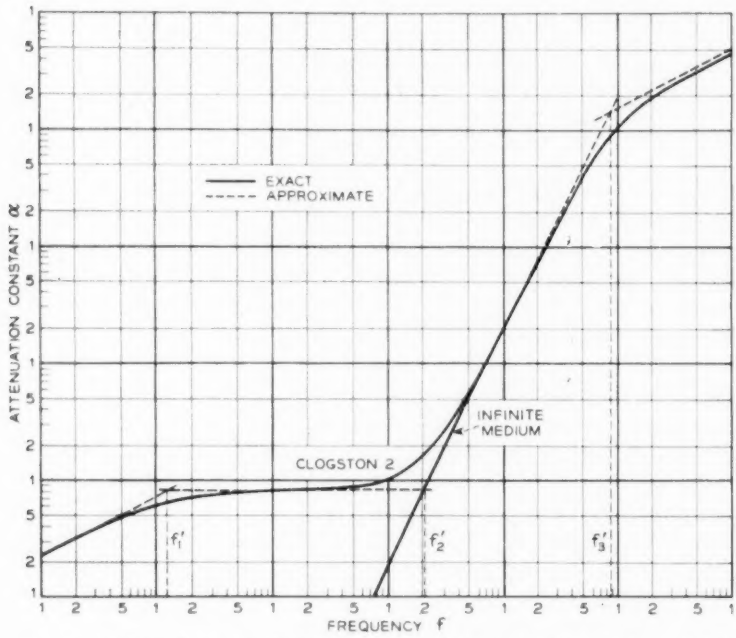


Fig. 20—Attenuation constants of a plane Clogston 2 line and an infinite laminated medium versus frequency on log-log scale.

numerical values are of no special significance, having been chosen solely for convenience in plotting.

So far as the practical applications of Clogston lines are concerned, we are primarily interested in the frequency range $f_1' \leqq f \leqq f_2'$, where the attenuation constant is essentially independent of frequency. To determine the rate at which the attenuation constant of the pth mode begins to deviate from its "flat" value as the frequency is increased, we write equation (459) in the form

$$\begin{aligned} \alpha &= \frac{p^2\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}\left[1 + \frac{4t_1^2T_1^2}{3p^2\pi^2\delta_1^4}\right] \\ &= \frac{p^2\pi^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}a^2}\left[1 + \frac{4t_1^2T_1^2\mu_1^2g_1^2f^2}{3p^2}\right]. \end{aligned} \tag{468}$$

The two terms in the square brackets are equal, and the attenuation constant is double its "flat" value, when $f = f_2'$, a result which is in very good agreement with the calculated values shown in Fig. 20.

The maximum permissible thickness of the conducting layers in a

plane Clogston 2 with high-impedance walls, if the attenuation constant of the pth mode is not to exceed its "flat" value α_0 by more than a specified small fraction $\Delta\alpha/\alpha_0$ at a top frequency f_m, is easily shown from (468) to be

$$t_1 = \frac{\sqrt{3}\ p}{2T_1\mu_1 g_1 f_m}\sqrt{\frac{\Delta\alpha}{\alpha_0}}. \tag{469}$$

Measuring f_m in Mc·sec^{-1} and thicknesses in mils, and putting in numerical values for copper, we obtain

$$(t_1)_{\text{mils}} = \frac{36.84p}{(2T_1)_{\text{mils}}(f_m)_{\text{Mc}}}\sqrt{\frac{\Delta\alpha}{\alpha_0}}. \tag{470}$$

We see from (461) that for fixed θ, α_0 is inversely proportional to $(a/p)^2$, where a is the total thickness of the stack and p is the mode number, while from (469) the permissible value of t_1 for a certain fractional change in attenuation is inversely proportional to (a/p), and also inversely proportional to the top frequency f_m.

It is interesting to compare equation (470) for the principal mode ($p = 1$) with equation (199) of Section V for the principal mode in an extreme Clogston 1 line with copper conducting layers. Since in a plane line $\Delta R/R_0 = \Delta\alpha/\alpha_0$, equation (199) may be written

$$(t_1)_{\text{mils}} = \frac{40.62}{(2T_1)_{\text{mils}}(f_m)_{\text{Mc}}}\sqrt{\frac{\Delta\alpha}{\alpha_0}}, \tag{471}$$

where $2T_1$ represents the total thickness of copper in both stacks. We expect that for partially filled plane Clogston lines with different proportions of the available space occupied by stacks, the maximum permissible layer thickness will be given by equations similar to (470) and (471), with values of the numerical coefficient intermediate between 36.84 and 40.62.

We turn next to a discussion of coaxial Clogston cables with finite laminae. A coaxial Clogston 2 is shown schematically in Fig. 11, and an enlarged view of part of the laminated stack in Fig. 4. The boundary conditions which apply to circular transverse magnetic waves on this structure are satisfied if at every point the sum of the radial impedances looking in opposite directions is zero. If we knew the explicit relation between the impedances at the two surfaces of the stack in terms of the stack parameters and the propagation constant γ, the impedance-matching conditions at the inner core and the outer sheath would yield a transcendental equation for the propagation constants of

the various possible modes. If the coaxial layers are of finite thickness, however, the relation between the surface impedances of the stack involves the product of as many different matrices as there are layers in the whole stack, and this matrix product is not suited to analytic treatment. We shall therefore approach the problem from another point of view.

We have seen that if the conducting layers in a laminated transmission line are sufficiently thin compared to the skin depth, the attenuation constant is essentially independent of frequency. In practice it is important to know how rapidly the attenuation constant of a Clogston cable with finite laminae begins to deviate from its low-frequency value as the frequency is increased. In accordance with the results for the parallel-plane line, we expect the initial increase to be proportional to the square of the frequency. We shall derive the term proportional to f^2 in the attenuation constant of the coaxial line on the basis of the following assumptions:

We assume that the macroscopic current distribution in a coaxial Clogston 2 is independent of frequency, and hence is given by the expressions which have already been derived for the case of infinitesimally thin laminae. (It is easy to show that this assumption is valid for a *plane* Clogston 2.) If the conducting layers are of finite thickness, then each carries a definite finite fraction of the total current in the line. At low frequencies the current density in any given layer is approximately uniform, but as the frequency is increased it becomes nonuniform because of the development of skin effect, and the power dissipated in the layer is increased. We shall calculate the total power dissipated in the stack, and the corresponding attenuation constant, up to terms in f^2.

Let the jth conducting layer in the stack be a hollow cylinder of conductivity g_1, inner radius ρ_{j-1}, and thickness t_1. Thus if there are $2n$ double layers we have $\rho_0 = a$ and $\rho_{2n} = b$, where as usual a and b denote the inner and outer radii of the whole stack. Let the total current flowing in the positive z-direction inside $\rho = \rho_{j-1}$ be I_{j-1}, and let the current flowing in the jth conducting layer be ΔI_j. It is shown in Appendix III that the average power dissipated per unit length in the jth conductor is approximately

$$\Delta P_j = \frac{1}{4\pi g_1 t_1 \rho_{j-1}} \left[|\Delta I_j|^2 + \frac{t_1^4}{3\delta_1^4} |I_{j-1}|^2 \right], \tag{472}$$

up to terms in $(t_1/\delta_1)^4$, where curvature corrections of the order of t_1/ρ_{j-1} have been neglected in comparison with unity. Presumably the only layers for which it may not be justifiable to neglect curvature cor-

rections will be the extreme inner layers, which occupy at most a small fraction of the total volume of the stack.

The average current density $\overline{J}_z$ in a Clogston 2 with infinitesimally thin laminae is given by equation (305) of Section VIII; namely, writing χ_p for the pth mode and dropping $e^{-\gamma z}$,

$$\overline{J}_z = H_0 \chi_p C_0(\chi_p \rho), \tag{473}$$

where H_0 is an arbitrary amplitude constant. For $n = 0$ and 1, $C_n(\chi_p \rho)$ denotes the combination of Bessel functions

$$C_n(\chi_p \rho) = N_1(\chi_p b) J_n(\chi_p \rho) - J_1(\chi_p b) N_n(\chi_p \rho); \tag{474}$$

and χ_p is the pth positive root of

$$C_1(\chi a) = N_1(\chi b) J_1(\chi a) - J_1(\chi b) N_1(\chi a) = 0. \tag{475}$$

According to equation (434) of Section X, we may write

$$\chi_p = \frac{p\pi f_p(a/b)}{b - a}, \tag{476}$$

where the functions $f_p(a/b)$ are of the order of unity. The total current $I(\rho)$ flowing in the positive z-direction between the inner core and a cylinder of arbitrary radius ρ is just

$$I(\rho) = 2\pi \int_a^\rho \rho \overline{J}_z \, d\rho = 2\pi H_0 \rho C_1(\chi_p \rho). \tag{477}$$

The thickness of the jth conducting layer in a stack of finite layers may be written

$$t_1 = \theta(t_1 + t_2) = \theta(\rho_j - \rho_{j-1}) = \theta \Delta\rho_j, \tag{478}$$

where $\Delta\rho_j$ represents the thickness $t_1 + t_2$ of the jth double layer. Hence approximately

$$\Delta I_j = 2\pi \rho_{j-1} \overline{J}_z \Delta\rho_j = 2\pi H_0 \chi_p \rho_{j-1} C_0(\chi_p \rho_{j-1}) \Delta\rho_j, \tag{479}$$

it being remembered that the conduction current in the conducting layer is essentially equal to the total current in the double layer, since the displacement currents are negligible. The current flowing inside the radius ρ_{j-1} is, from (477),

$$I_{j-1} = 2\pi H_0 \rho_{j-1} C_1(\chi_p \rho_{j-1}), \tag{480}$$

and so the power dissipated per unit length in the jth conductor is

given by (472) as

$$\Delta P_j = \frac{\pi H_0 H_0^* \rho_{j-1}}{\theta q_1} \left[\chi_p^2 C_0'^2(\chi_p \rho_{j-1}) + \frac{\theta^2 l_1^2}{3\delta_1^4} C_1'^2(\chi_p \rho_{j-1}) \right] \Delta\rho_j . \quad (481)$$

The total power ΔP dissipated per unit length in the whole stack is obtained by summing ΔP_j over j. Approximately the sum by an integral, we have

$$\begin{aligned} \Delta P &= \frac{\pi H_0 H_0^*}{\bar{g}} \int_a^b \rho[\chi_p^2 C_0'^2(\chi_p \rho) + \frac{\theta^2 l_1^2}{3\delta_1^4} C_1'^2(\chi_p \rho)]\, d\rho \\ &= \frac{\pi H_0 H_0^* \chi_p^2}{2\bar{g}} \left[1 + \frac{\theta^2 l_1^2}{3\chi_p^2 \delta_1^4} \right] [b^2 C_0'^2(\chi_p b) - a^2 C_0'^2(\chi_p a)]. \end{aligned} \quad (482)$$

The average transmitted power P when the laminae are infinitesimally thin is

$$\begin{aligned} P &= \text{Re } \tfrac{1}{2} \int_0^{2\pi} \int_a^b \bar{E}_\rho H_\phi^* \rho\, d\rho\, d\phi \\ &= \pi \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 H_0^* \int_a^b \rho C_1'^2(\chi_p \rho)\, d\rho \\ &= \tfrac{1}{2}\pi \sqrt{\frac{\bar{\mu}}{\bar{\epsilon}}}\, H_0 H_0^* [b^2 C_0'^2(\chi_p b) - a^2 C_0'^2(\chi_p a)]. \end{aligned} \quad (483)$$

If we assume the same value for P when the laminae are of finite thickness, then from (482) and (483) the attenuation constant of the line is

$$\alpha = \frac{\Delta P}{2P} = \frac{\chi_p^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g}} \left[1 + \frac{\theta^2 l_1^2}{3\chi_p^2 \delta_1^4} \right]. \quad (484)$$

The similarity of equation (484) to equation (468) for the parallel-plane line becomes obvious if we write χ_p in the form (476) and denote the total thickness $\theta(b - a)$ of conducting material in the coaxial stack by $2T_1$. We then have

$$\begin{aligned} \alpha &= \frac{p^2 \pi^2 f_p^2(a/b)}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g}(b-a)^2} \left[1 + \frac{4 l_1^2 T_1^2}{3 p^2 \pi^2 f_p^2(a/b) \delta_1^4} \right] \\ &= \frac{p^2 \pi^2 f_p^2(a/b)}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g}(b-a)^2} \left[1 + \frac{4 l_1^2 T_1^2 \mu_1^2 g_1^2 f^2}{3 p^2 f_p^2(a/b)} \right], \end{aligned} \quad (485)$$

and as the ratio a/b approaches unity the function $f_p^2(a/b)$ approaches unity and (485) becomes identical with (468). We recall that $f_1^2(a/b)$ was plotted against a/b in Fig. 12. For the principal mode in a cable

with no inner core ($a = 0$), equation (485) takes the form

$$\alpha = \frac{7.341}{\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g} b^2} [1 + 0.8963 \bar{t}_1^2 T_1^2 \mu_1^2 g_1^2 f^2]. \tag{486}$$

It should be emphasized that whereas equation (468) was obtained from a rigorous solution of the boundary-value problem for the plane line, equation (485) for the coaxial cable has been derived on the basis of certain physical assumptions and approximations whose effect on the accuracy of the final result is not very easy to estimate. Presumably one might check the accuracy of (485) for a particular Clogston cable by setting up the matrix relation between the known surface impedances of the core and the outer sheath and solving numerically for the propagation constant. It should not be too difficult to solve the matrix equation by cut-and-try methods for a cable having, say, two hundred double layers, if the matrix of each double layer were assumed to be given by equations (88) of Section III, and high-speed computing machinery were used to perform the matrix multiplications. In the absence of any such numerical results, however, we shall merely assume that equation (485) furnishes a reasonable approximation to the attenuation constant of a coaxial Clogston 2 in the frequency range $f_1' \leqq f \leqq f_3'$, where f_1' and f_3' are the critical frequencies defined by (467).

The first conclusion which we can draw from (485) is that the maximum permissible thickness of the conducting layers in a coaxial Clogston 2 with high-impedance boundaries, if the attenuation constant of the pth mode is not to exceed its "flat" value α_0 by more than a specified small fraction $\Delta\alpha/\alpha_0$ at a top frequency f_m, is

$$t_1 = \frac{\sqrt{3}\; p f_p(a/b)}{2 T_1 \mu_1 g_1 f_m} \sqrt{\frac{\Delta\alpha}{\alpha_0}}; \tag{487}$$

or, putting in numerical values for copper,

$$(t_1)_{\text{mils}} = \frac{36.84 p f_p(a/b)}{(2T_1)_{\text{mils}} (f_m)_{\text{Mc}}} \sqrt{\frac{\Delta\alpha}{\alpha_0}}. \tag{488}$$

For the principal mode in a Clogston cable with no inner core, this becomes

$$(t_1)_{\text{mils}} = \frac{44.93}{(2T_1)_{\text{mils}} (f_m)_{\text{Mc}}} \sqrt{\frac{\Delta\alpha}{\alpha_0}}. \tag{489}$$

As a second application of equation (485), we shall determine the upper crossover frequency at which the attenuation constant of a Clogston 2 is equal to the attenuation constant of a conventional coaxial

cable of the same size. Since the lower crossover frequency was found at the end of Section VIII, we shall then know the theoretical limits of the frequency range over which a given Clogston cable can have lower loss than the corresponding standard coaxial.

According to equation (317) of Section VIII, a conventional coaxial cable of radius b and optimum proportions has an attenuation constant

$$\alpha = \frac{1.796}{\sqrt{\mu_0/\epsilon_0}\, g_1 \delta_1 b} = \frac{1.796\sqrt{\pi\mu_1 g_1 f}}{\sqrt{\mu_0/\epsilon_0}\, g_1 b}. \tag{490}$$

We shall assume that the upper crossover occurs in the high-frequency range where the attenuation constant of a Clogston 2 is approximately proportional to f^2. Then for the pth mode in a cable with no inner core ($a = 0$), equation (485) gives

$$\alpha = \frac{2\pi^2 t_1^2 T_1^2 \mu_1^2 g_1^2 f^2}{3\sqrt{\bar{\mu}/\bar{\epsilon}}\, \bar{g} b^2} = \frac{\pi^2 t_1^2 \mu_1^2 \bar{g} f^2}{6\sqrt{\bar{\mu}/\bar{\epsilon}}}. \tag{491}$$

A little algebra shows that the two attenuation constants are equal when

$$f = \frac{1}{\pi\mu_1 g_1}\left[\frac{10.77}{2T_1 t_1^2}\sqrt{\frac{\bar{\mu}\epsilon_0}{\mu_0\bar{\epsilon}}}\right]^3. \tag{492}$$

If the conventional cable is air-filled, then assuming copper conductors and no magnetic materials, we find that equation (492) becomes, numerically,

$$f_{\text{Mc}} = 33.02\left[\frac{1}{(2T_1)_{\text{mils}}(t_1^2)_{\text{mils}}}\sqrt{\frac{1-\theta}{\epsilon_{2r}}}\right]^3. \tag{493}$$

If we consider a 3/8-inch Clogston cable with 0.1-mil copper conductors, 0.05-mil polyethylene insulators, and no inner core, then

$$\begin{aligned} b &= 187.5 \text{ mils} & \theta &= 2/3 \\ 2T_1 &= 125 \quad \text{mils} & \epsilon_{2r} &= 2.26 \\ t_1 &= 0.1 \quad \text{mils} \end{aligned} \tag{494}$$

We found in Section VIII that the lower crossover frequency for this cable is about 50 kc·sec^{-1}, while from equation (493) the upper crossover frequency turns out to be 15 Mc·sec^{-1}.

We next discuss the problem of maximizing the frequency band over which the attenuation constant of a Clogston cable of given diameter

does not exceed a specified value.[26] We suppose that the thickness t_1 of the conductors is fixed, but that the proportion of conducting material in the cable may be adjusted by changing the thickness of the insulators. Let α_m be the value of the attenuation constant which is not to be exceeded in the operating frequency range, and let f_m be the frequency at which this maximum attenuation is reached. What should be the fraction θ of conducting material in the cable in order to maximize f_m? It is tacitly assumed that α_m is at least slightly greater than the minimum "flat" attenuation constant which is possible with a cable of the given diameter, since obviously we do not wish to work entirely in the very-low-frequency range.

In the frequency range of interest the attenuation constant of the pth mode is given by equation (484), which may be written

$$\alpha = \frac{\chi_p^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}} + \frac{\theta^2 t_1^2 \pi^2 \mu_1^2 g_1^2 f^2}{6\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}, \tag{495}$$

where χ_p is a root of (475) and independent of θ. Solving (495) for the frequency f_m at which α is equal to α_m, and substituting for $\bar{\epsilon}$, $\bar{\mu}$, and $\bar{g}$ from (268), we obtain

$$f_m = \frac{\sqrt{3}}{\pi \mu_1 g_1 t_1}\left[\frac{2[\theta\mu_1 + (1-\theta)\mu_2]^{\frac{1}{2}}[(1-\theta)/\epsilon_2]^{\frac{1}{2}} g_1 \alpha_m}{\theta} - \frac{\chi_p^2}{\theta^2}\right]^{\frac{1}{2}}. \tag{496}$$

A routine calculation shows that f_m is a maximum, considered as a function of θ, when θ satisfies

$$\frac{\alpha_m g_1 \theta[\theta\mu_1 + 2(1-\theta)\mu_2]}{[\theta\mu_1 + (1-\theta)\mu_2]^{\frac{1}{2}}(1-\theta)^{\frac{1}{2}}\epsilon_2^{\frac{1}{2}}} = 2\chi_p^2. \tag{497}$$

Equation (497) is easily reduced to a quartic equation in θ, which may be solved without difficulty when the other parameters are given. The maximum value of f_m is then obtained by substituting θ back into (496).

We shall now investigate in more detail the case in which

$$\mu_1 = \mu_2, \tag{498}$$

that is, the permeabilities of the conducting and insulating layers are equal. In this case the low-frequency attenuation constant α_0, which is just the first term on the right side of equation (495), is given by

$$\alpha_0 = \frac{\chi_p^2}{2\theta\sqrt{1-\theta}\,\sqrt{\mu_2/\epsilon_2}\,g_1}, \tag{499}$$

[26] A similar problem was first investigated in an unpublished memorandum by H. S. Black.

and α_0 has a minimum when

$$\theta = 2/3. \tag{500}$$

The minimum value of the low-frequency attenuation constant, which we may call α_{00}, is just

$$\alpha_{00} = \frac{3\sqrt{3}\ \chi_p^2}{4\sqrt{\mu_2/\epsilon_2}\ g_1}. \tag{501}$$

Writing

$$\chi_p^2 = \frac{4\sqrt{\mu_2/\epsilon_2}\ g_1\alpha_{00}}{3\sqrt{3}}, \tag{502}$$

we find that equation (496) takes the form

$$f_m = \frac{\sqrt{3}\ \chi_p}{\pi\mu_1 g_1 l_1}\left[\frac{3\sqrt{3}\ (1-\theta)^{\frac{1}{2}}}{2\theta}\frac{\alpha_m}{\alpha_{00}} - \frac{1}{\theta^2}\right]^{\frac{1}{2}}, \tag{503}$$

for any value of θ. From equation (497), f_m is a maximum when θ satisfies

$$\frac{\theta(2-\theta)}{(1-\theta)^{\frac{1}{2}}} = \frac{8}{3\sqrt{3}}\frac{\alpha_{00}}{\alpha_m}, \tag{504}$$

which is equivalent to the quartic equation

$$\theta^4 - 4\theta^3 + 4\theta^2 + \frac{64\alpha_{00}^2}{27\alpha_m^2}(\theta - 1) = 0. \tag{505}$$

If θ_m is the root of (505) which lies between zero and one, then the corresponding value of f_m is

$$f_m = \frac{\sqrt{3}\ \chi_p}{\pi\mu_1 g_1 l_1 \theta_m}\left[\frac{2-3\theta_m}{2-\theta_m}\right]^{\frac{1}{2}}. \tag{506}$$

We observe from either (503) or (506) that f_m is inversely proportional to l_1.

The values of θ_m and $\theta_m^{-1}[(2-3\theta_m)/(2-\theta_m)]^{\frac{1}{2}}$ are plotted in Fig. 21 against α_m/α_{00}, which is just the ratio of the maximum attenuation constant to the minimum low-frequency attenuation constant which can be achieved with a Clogston cable of the same diameter. When α_m/α_{00} is unity, then $\theta_m = 2/3$ and f_m is zero to the present approximation (a better estimate of f_m would be the critical frequency f_1' defined by equation (467)). For values of α_m/α_{00} greater than about five, θ_m is

given to a good approximation by

$$\theta_m \approx \frac{4\alpha_{00}}{3\sqrt{3}\,\alpha_m} \approx \frac{0.77\alpha_{00}}{\alpha_m}, \tag{507}$$

while

$$f_m \approx \frac{2.25\chi_p}{\pi\mu_1 g_1 t_1}\frac{\alpha_m}{\alpha_{00}}. \tag{508}$$

The low-frequency attenuation constant α_0 of a Clogston cable with $\theta = \theta_m$ will of course be greater than α_{00} if θ_m is not equal to 2/3. This is not really a disadvantage, however, since by assumption we only wish to insure that $\alpha \leqq \alpha_m$ over the operating band, and the nearer α approaches to α_m over the whole band the less serious will be the equalization problem. It may be shown that the ratio α_0/α_m decreases from unity toward one-half as α_m/α_{00} is increased indefinitely. Physically this means that the low-frequency attenuation constant of an optimum Clogston cable is always at least half as great as the attenuation constant at the upper end of the band, and the cable never contains more conducting material than would correspond to a total stack thickness of about two effective skin depths at the highest operating frequency.

We conclude with a few numerical formulas relating to the principal mode in a completely filled Clogston cable with copper conductors and no inner core. The low-frequency attenuation constant α_0 of such a

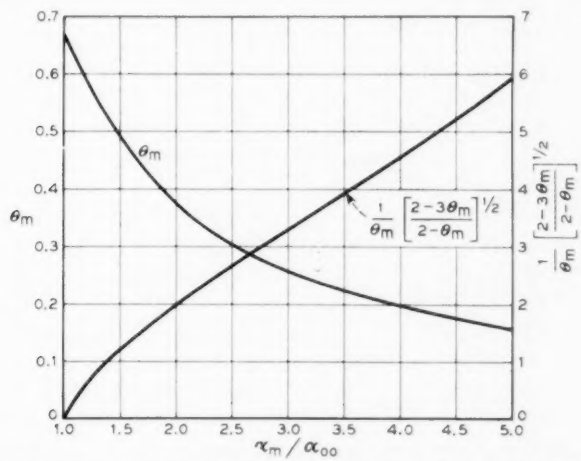


Fig. 21—Curves related to the optimum fraction θ_m of conducting material in Clogston cables with finite laminae, as a function of the attenuation ratio α_m/α_{00}.

cable is given by

$$\begin{aligned}\alpha_0 &= \frac{0.521\sqrt{\epsilon_{2r}}}{\theta\sqrt{1-\theta}\,b_{\text{mils}}^2}\ \text{nepers}\cdot\text{meter}^{-1} \\ &= \frac{0.728\times 10^4\sqrt{\epsilon_{2r}}}{\theta\sqrt{1-\theta}\,b_{\text{mils}}^2}\ \text{db}\cdot\text{mile}^{-1},\end{aligned} \tag{509}$$

for any value of θ, while if $\theta = 2/3$, we have

$$\begin{aligned}\alpha_{00} &= \frac{1.353\sqrt{\epsilon_{2r}}}{b_{\text{mils}}^2}\ \text{nepers}\cdot\text{meter}^{-1} \\ &= \frac{1.891\times 10^4\sqrt{\epsilon_{2r}}}{b_{\text{mils}}^2}\ \text{db}\cdot\text{mile}^{-1}.\end{aligned} \tag{510}$$

The frequency f_m as a function of the ratio α_m/α_{00} is

$$(f_m)_{\text{Mc}} = \frac{44.93}{(t_1)_{\text{mils}}b_{\text{mils}}}\left[2.598\,\frac{(1-\theta)^{\frac{1}{2}}}{\theta}\,\frac{\alpha_m}{\alpha_{00}} - \frac{1}{\theta^2}\right]^{\frac{1}{2}}, \tag{511}$$

for any θ, and when $\theta = \theta_m$ the expression for f_m becomes

$$(f_m)_{\text{Mc}} = \frac{44.93}{(t_1)_{\text{mils}}b_{\text{mils}}}\,\frac{1}{\theta_m}\left[\frac{2-3\theta_m}{2-\theta_m}\right]^{\frac{1}{2}}, \tag{512}$$

where the factor involving θ_m is plotted against α_m/α_{00} in Fig. 21.

If we consider a 3/8-inch Clogston cable with 0.1-mil copper conductors and polyethylene insulation ($\epsilon_{2r} = 2.26$), we find

$$\alpha_{00} = 0.809\ \text{db}\cdot\text{mile}^{-1}. \tag{513}$$

If we set

$$\alpha_m = 2\alpha_{00} = 1.618\ \text{db}\cdot\text{mile}^{-1}, \tag{514}$$

then it turns out that

$$\theta_m = 0.3745, \tag{515}$$

so that the insulating layers should be 0.167 mil thick. The low-frequency attenuation constant for θ_m is

$$\alpha_0 = 1.300\alpha_{00} = 1.051\ \text{db}\cdot\text{mile}^{-1}, \tag{516}$$

and α_m is reached at a frequency

$$f_m = 4.70\ \text{Mc}\cdot\text{sec}^{-1}. \tag{517}$$

If we had used $\theta = 2/3$, we should have reached α_m at a frequency of

3.59 Mc·sec^{-1}, which is only 76.5 per cent of the frequency given by (517). An ordinary air-filled coaxial cable of the same size would have an attenuation constant equal to α_{00} at about 50 kc·sec^{-1} and equal to α_m ($= 2\alpha_{00}$) at about 200 kc·sec^{-1}.

It should be borne in mind that in the preceding example we have neglected the effects of dielectric loss and of stack nonuniformity. Neither of these effects can be completely eliminated in a physical Clogston cable, and both will exert increasingly adverse influences on the attenuation constant as the frequency is raised.

XII. EFFECT OF NONUNIFORMITY OF LAMINATED MEDIUM

In the previous analysis of laminated transmission lines we have treated only perfectly uniform structures, in which every conducting layer is identical to every other conducting layer in thickness and in electrical properties, and all the insulating layers are similarly identical to each other. In practice, however, it will not be possible to lay down large numbers of absolutely identical thin layers, and we therefore need to know the effect on transmission of slight nonuniformities in the laminated stacks. Some indication that stack uniformity will be a very critical problem in laminated cables which are expected to give large improvements in attenuation over conventional coaxial cables of the same size may be obtained from the results of Section VI, which showed that in a Clogston 1 line, where the phase velocity is determined by the $\mu\epsilon$ product of the main dielectric, this product must be controlled very accurately to maintain the desired deep penetration of current into the laminated stacks. In a Clogston 2, where the main dielectric has been replaced by extensions of the stacks, one might expect similarly stringent requirements on the uniformity of the laminated material if the desired current distribution is to be maintained.

In this section we estimate the effects of stack nonuniformity by studying some particular idealized cases of nonuniformity in a parallel-plane Clogston 2 with infinitesimally thin layers, in which the average electrical properties of the stack vary only in the direction perpendicular to the layers. The principal conclusion is that if one attempts to realize with a Clogston line an attenuation constant which is a small fraction, say of the order of one-tenth, of the attenuation constant of a conventional line of the same dimensions at the same frequency, then long-range variations in the properties of the stack (as distinguished from short-range random fluctuations) must be controlled to within a few parts in 10,000. The price is less steep if the overall improvement sought

is less, but in all practical cases it appears that the average properties of the stack must be held constant against slow variations to a fraction of a per cent. The requirement of extraordinarily high precision is in addition to the requirement that the individual layers must be extremely thin if a Clogston cable is to improve on a conventional coaxial cable at all in the megacycle frequency range.

For purposes of analysis, we consider a parallel-plane Clogston 2 transmission line bounded by infinite-impedance sheaths at $y = \pm \frac{1}{2}a$, as shown schematically in Fig. 10. The individual layers are supposed to be infinitesimally thin, so that near any given point the average electrical constants of the stack are

$$\begin{aligned} \bar{\epsilon} &= \epsilon_2/(1 - \theta), \\ \bar{\mu} &= \theta\mu_1 + (1 - \theta)\mu_2 , \\ \bar{g} &= \theta g_1 . \end{aligned} \tag{518}$$

The quantities $\bar{\epsilon}$, $\bar{\mu}$, and $\bar{g}$ may vary, continuously or with a finite number of finite discontinuities, as functions of the transverse coordinate y, owing to variations in any or all of μ_1 , g_1 , μ_2 , ϵ_2 , and θ; but they are not supposed to vary with x or z.

We shall be concerned with modes in which the fields are independent of x, and in which the only field components are H_x , $\bar{E}_y$, and E_z . Then Maxwell's equations are given by (269) of Section VIII, and reduce, if we write the field components in the form $H_x(y)e^{-\gamma z}$, $\bar{E}_y(y)e^{-\gamma z}$, and $E_z(y)e^{-\gamma z}$, to

$$\begin{aligned} -\gamma H_x &= i\omega\bar{\epsilon}\bar{E}_y , \\ dH_x/dy &= -\bar{g}E_z , \\ -\gamma\bar{E}_y - dE_z/dy &= i\omega\bar{\mu}H_x . \end{aligned} \tag{519}$$

If we eliminate $\bar{E}_y$ and E_z from these equations we obtain

$$\frac{d^2H_x}{dy^2} - \frac{1}{\bar{g}}\frac{d\bar{g}}{dy}\frac{dH_x}{dy} - i\omega\bar{\mu}\bar{g}\left[1 + \frac{\gamma^2}{\omega^2\bar{\mu}\bar{\epsilon}}\right]H_x = 0, \tag{520}$$

where H_x and E_z must be continuous at any points of discontinuity of $\bar{\epsilon}$, $\bar{\mu}$, or $\bar{g}$. The tangential magnetic field must vanish on the infinite-impedance surfaces at $y = \pm\frac{1}{2}a$; hence we have the boundary conditions

$$H_x(-\tfrac{1}{2}a) = H_x(\tfrac{1}{2}a) = 0. \tag{521}$$

These boundary conditions, taken in conjunction with the differential

equation (520), define the values of γ which are the propagation constants of the various modes of the line.

While it is possible to find special forms of the functions $\bar{\epsilon}(y)$, $\bar{\mu}(y)$, and $\bar{g}(y)$ such that (520) can be solved exactly in terms of known functions, it is easier to make certain approximations in the beginning which retain only the important terms. For this purpose we shall write

$$\begin{aligned} \bar{\epsilon} &= \bar{\epsilon}_0 + \Delta\bar{\epsilon}, \\ \bar{\mu} &= \bar{\mu}_0 + \Delta\bar{\mu}, \\ \bar{g} &= \bar{g}_0 + \Delta\bar{g}, \end{aligned} \tag{522}$$

where $\bar{\epsilon}_0$, $\bar{\mu}_0$, and $\bar{g}_0$ are constants representing the average values of $\bar{\epsilon}$, $\bar{\mu}$, and $\bar{g}$ across the stack, so that the average values of $\Delta\bar{\epsilon}$, $\Delta\bar{\mu}$, and $\Delta\bar{g}$ across the stack are zero.* Furthermore the fractional variations in the stack parameters will be assumed small compared to unity; in practical cases they will never be larger than a few per cent and will usually be only a fraction of one per cent.

Referring now to equation (520), we see that the coefficient of H_x contains the large factor $\omega\bar{\mu}\bar{g}$, which is of the order of $1/\delta_1^2$, as compared with the term d^2H_x/dy^2, which is presumably of the order of $(1/a^2)H_x$. Hence small changes in $\bar{\epsilon}$ and $\bar{\mu}$ will make relatively large changes in the coefficient of H_x, since γ^2 is a constant. On the other hand, the coefficient of dH_x/dy will be small for any reasonable variations in the small quantity $\Delta\bar{g}/\bar{g}_0$. Hence we shall neglect this term entirely and deal with the approximate equation

$$\frac{d^2H_x}{dy^2} - i\omega\bar{\mu}\bar{g}\left[1 + \frac{\gamma^2}{\omega^2\bar{\mu}\bar{\epsilon}}\right]H_x = 0. \tag{523}$$

If we substitute (522) into (523) and drop second order terms in $\Delta\bar{\epsilon}/\bar{\epsilon}_0$, $\Delta\bar{\mu}/\bar{\mu}_0$, and $\Delta\bar{g}/\bar{g}_0$, we find that the coefficient of H_x becomes

$$\begin{aligned} &-\frac{i\bar{g}}{\omega\bar{\epsilon}}\left[\omega^2\bar{\mu}\bar{\epsilon} + \gamma^2\right] \\ &\approx -\frac{i\bar{g}_0}{\omega\bar{\epsilon}_0}\left[1 + \frac{\Delta\bar{g}}{\bar{g}_0} - \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right]\left[\gamma^2 + \omega^2\bar{\mu}_0\bar{\epsilon}_0 + \omega^2\bar{\mu}_0\bar{\epsilon}_0\left(\frac{\Delta\bar{\mu}}{\bar{\mu}_0} + \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right)\right]; \end{aligned} \tag{524}$$

and if

$$\Gamma_t^2 = \frac{i\bar{g}_0}{\omega\bar{\epsilon}_0}\left[\omega^2\bar{\mu}_0\bar{\epsilon}_0 + \gamma^2\right], \tag{525}$$

* The present use of zero subscripts on $\bar{\epsilon}_0$, $\bar{\mu}_0$, and $\bar{g}_0$ has of course nothing to do with the earlier convention that associated zero subscripts with the main dielectric in Clogston I lines.

so that

$$\gamma^2 = -\omega^2\bar{\mu}_0\bar{\epsilon}_0 - (i\omega\bar{\epsilon}_0/\bar{g}_0)\Gamma_\ell^2, \tag{526}$$

then (524) becomes, approximately,

$$\begin{aligned} &-\left[1 + \frac{\Delta\bar{g}}{\bar{g}_0} - \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right]\left[\Gamma_\ell^2 + i\omega\bar{\mu}_0\bar{g}_0\left(\frac{\Delta\bar{\mu}}{\bar{\mu}_0} + \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right)\right] \\ &\approx -\left[\Gamma_\ell^2 + i\omega\bar{\mu}_0\bar{g}_0\left(\frac{\Delta\bar{\mu}}{\bar{\mu}_0} + \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right)\right]. \end{aligned} \tag{527}$$

In all cases of interest we shall find that $(\Delta\bar{\mu}/\bar{\mu}_0 + \Delta\bar{\epsilon}/\bar{\epsilon}_0)$ is smaller than or at most of the same order of magnitude as $\Gamma_\ell^2/i\omega\bar{\mu}_0\bar{g}_0$. Hence the differential equation (523) takes the approximate form

$$\frac{d^2H_x}{dy^2} - \left[\Gamma_\ell^2 + i\omega\bar{\mu}_0\bar{g}_0\left(\frac{\Delta\bar{\mu}}{\bar{\mu}_0} + \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0}\right)\right]H_x = 0, \tag{528}$$

where Γ_ℓ^2 is determined by the two-point boundary conditions (521).

The variations of the stack parameters appear in (528) only in the term $(\Delta\bar{\mu}/\bar{\mu}_0 + \Delta\bar{\epsilon}/\bar{\epsilon}_0)$, which is some as yet unspecified function of y. For convenience we shall write this term in the form

$$\frac{\Delta\bar{\mu}}{\bar{\mu}_0} + \frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0} = \frac{C}{\omega\bar{\mu}_0\bar{g}_0a^2}\varphi(y), \tag{529}$$

where C is a dimensionless parameter and $\varphi(y)$ is a function whose average value over the stack is zero, and whose maximum absolute value will usually be of the order of unity. It is worth noting that if the conducting and insulating layers all have equal permeabilities, then (529) becomes

$$\frac{\Delta\bar{\epsilon}}{\bar{\epsilon}_0} = \frac{C\delta_1^2}{2a^2\theta_0}\varphi(y), \tag{530}$$

where δ_1 is the skin depth in the average conducting layer and θ_0 is the average fraction of space filled with conducting material. If we solve the differential equation for different values of the scale factor C but the same $\varphi(y)$, we can calculate the effect of stack nonuniformities of the same type but different amplitudes, or the effect of nonuniformity in the same stack at different frequencies. In the latter case C is directly proportional to the frequency.

The final step in the transformation of the differential equation (528) will be to reduce it to dimensionless form by the substitutions

$$\begin{aligned} \xi &= y/a + \tfrac{1}{2}, \\ w(\xi) &= H_x(y), \\ f(\xi) &= \varphi(y), \\ \Lambda &= -\Gamma_\ell^2 a^2. \end{aligned} \tag{531}$$

Then on making use of (529) we get

$$d^2w/d\xi^2 + [\Lambda - iCf(\xi)]w(\xi) = 0, \tag{532}$$

with the boundary conditions

$$w(0) = w(1) = 0. \tag{533}$$

Once Λ has been determined for a particular mode, the propagation constant γ is obtained from (526) and (531), namely

$$\gamma = i\omega\sqrt{\bar{\mu}_0\bar{\epsilon}_0}\,[1 + \Lambda/i\omega\bar{\mu}_0\bar{g}_0a^2]^{\frac{1}{2}}. \tag{534}$$

Assuming as usual that the attenuation per radian is small, we find that the attenuation and phase constants are given by

$$\alpha = \operatorname{Re}\gamma = \operatorname{Re}\frac{\Lambda}{2\sqrt{\bar{\mu}_0/\bar{\epsilon}_0}\,\bar{g}_0a^2}, \tag{535}$$

$$\beta = \operatorname{Im}\gamma = \omega\sqrt{\bar{\mu}_0\,\bar{\epsilon}_0} + \operatorname{Im}\frac{\Lambda}{2\sqrt{\bar{\mu}_0/\bar{\epsilon}_0}\,\bar{g}_0a^2}. \tag{536}$$

The eigenvalues Λ of the differential equation (532) with boundary conditions (533) may be found analytically for some simple forms of $f(\xi)$, or numerically using a differential analyzer for any given $f(\xi)$ which does not fluctuate too rapidly. When $C = 0$, as in the case of a perfectly uniform stack, the eigenvalues are obviously

$$\Lambda_1 = \pi^2, \qquad \Lambda_2 = 4\pi^2, \cdots, \tag{537}$$

corresponding to the eigenfunctions

$$w_1 = \sin \pi\xi, \qquad w_2 = \sin 2\pi\xi, \cdots. \tag{538}$$

As C varies continuously, we expect the eigenvalues and eigenfunctions to vary continuously in a manner depending on $f(\xi)$. In the following paragraphs we shall discuss the behavior of Λ_1, and sometimes also Λ_2, as a function of C for various simple types of nonuniformity.

(*i*) $f(\xi)$ *constant except at single discontinuity.* Let

$$f(\xi) = \begin{cases} -\dfrac{1}{2\xi_0}, & 0 \leqq \xi < \xi_0, \\ \dfrac{1}{2(1-\xi_0)}, & \xi_0 < \xi \leqq 1, \end{cases} \tag{539}$$

where ξ_0 is some fixed number between 0 and 1 but not, in the cases of interest, extremely close to either 0 or 1. Solutions of (532) satisfying the boundary conditions (533) are obviously

$$\begin{aligned} w(\xi) &= A \sin [\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi, & 0 \leqq \xi < \xi_0, \\ w(\xi) &= B \sin [\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}}(1-\xi), & \xi_0 < \xi \leqq 1, \end{aligned} \tag{540}$$

where A and B are arbitrary constants. The requirements that w and $dw/d\xi$ be continuous* at $\xi = \xi_0$ lead to the equations

$$\begin{aligned} &A \sin [\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0 = B \sin [\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}}(1-\xi_0), \\ &A[\Lambda + iC/2\xi_0]^{\frac{1}{2}} \cos [\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0 \\ &\quad = -B[\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}} \cos [\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}}(1-\xi_0), \end{aligned} \tag{541}$$

which will be consistent if this characteristic equation is satisfied:

$$\frac{\tan [\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0}{[\Lambda + iC/2\xi_0]^{\frac{1}{2}}} + \frac{\tan [\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}}(1-\xi_0)}{[\Lambda - iC/2(1-\xi_0)]^{\frac{1}{2}}} = 0. \tag{542}$$

The roots in Λ of equation (542) are the eigenvalues of the problem; the eigenfunction corresponding to any given eigenvalue is given by equations (540) after the ratio B/A is determined from either of equations (541).

It is easy to show that when $C = 0$, the roots of (542) are $\Lambda_1 = \pi^2$, $\Lambda_2 = 4\pi^2, \cdots$. For large values of C, representing relatively great differences between the two parts of the stack, physical considerations lead us to expect that there will be pairs of modes, one member of each pair being essentially confined to each part of the stack and having a propagation constant determined approximately by the width of that part. It may in fact be shown that the asymptotic expression for the eigenvalue of the mode which is essentially confined to the region $0 \leqq \xi < \xi_0$ is

$$\Lambda \approx \frac{\pi^2}{\xi_0^2}\left[1 - 2\sqrt{\frac{(1-\xi_0)}{\xi_0 C}}\right] - i\left[\frac{C}{2\xi_0} - \frac{2\pi^2}{\xi_0^2}\sqrt{\frac{(1-\xi_0)}{\xi_0 C}}\right]; \tag{543}$$

* The continuity of $dw/d\xi$ is a consequence of the continuity of E_z, provided that we neglect any discontinuity in g at $\xi = \xi_0$.

and the asymptotic expression for the eigenvalue of the mode which is essentially confined to $\xi_0 < \xi \leqq 1$ is

$$\Lambda \approx \frac{\pi^2}{(1-\xi_0)^2}\left[1 - 2\sqrt{\frac{\xi_0}{(1-\xi_0)C}}\right] + i\left[\frac{C}{2(1-\xi_0)} - \frac{2\pi^2}{(1-\xi_0)^2}\sqrt{\frac{\xi_0}{(1-\xi_0)C}}\right]. \tag{544}$$

It is clear that if $\xi_0 < \frac{1}{2}$ the latter mode has the smaller attenuation constant, while if $\xi_0 > \frac{1}{2}$ the former mode has the smaller attenuation constant.

It is not difficult, although the details will be omitted here, to investigate the behavior of the eigenvalues for small C and to show that no matter whether $\xi_0 < \frac{1}{2}$ or $\xi_0 > \frac{1}{2}$, the eigenvalue which starts from π^2 at $C = 0$ tends to the asymptotic value which has the smaller real part, so that this eigenvalue, whether its asymptotic form be given by (543) or (544), may be called Λ_1. It appears that if $\xi_0 < \frac{1}{2}$, then Im Λ_1 is positive for positive C, while if $\xi_0 > \frac{1}{2}$, then Im Λ_1 is negative for positive C.

An interesting mathematical phenomenon appears when $\xi_0 = \frac{1}{2}$, so that the discontinuity in $f(\xi)$ is exactly at the center of the stack. In this case, when C is small Λ_1 and Λ_2 are both real, Λ_1 being somewhat greater than π^2 and Λ_2 somewhat less than $4\pi^2$. For a certain value of C the two eigenvalues coincide; this value is approximately

$$C = 17.9, \qquad \Lambda_1 = \Lambda_2 = 25.6. \tag{545}$$

For larger values of C, Λ_1 and Λ_2 are complex conjugates (it seems to be immaterial which is which) whose asymptotic forms are given by (543) and (544) with $\xi_0 = \frac{1}{2}$.

Approximate values of Λ_1 and Λ_2 were found for the symmetric case, $\xi_0 = 0.5$, and for one unsymmetric case, $\xi_0 = 0.6$, on the Laboratories' general purpose analog computer for $0 \leqq C \leqq 100$, and were refined afterward by desk computation, using a method of successive approximations to solve equation (542). The real and imaginary parts of Λ_1/π^2 and Λ_2/π^2 are plotted in Fig. 22 for the symmetric case, where it should be noted that different vertical scales are used for Re Λ/π^2 and Im Λ/π^2. The corresponding eigenfunctions $w_1(\xi)$ and $w_2(\xi)$ are shown in Fig. 23 for $C = 0$, $C = 17.9$, which corresponds to equal eigenvalues, and $C = 100$. It will be recalled that $w(\xi)$ is equal to $H_x(y)$, and the other field components can be derived from H_x by equations (519) if desired. Fig. 24 shows plots of Λ_1/π^2 and Λ_2/π^2 for the unsymmetric case $\xi_0 = 0.6$.

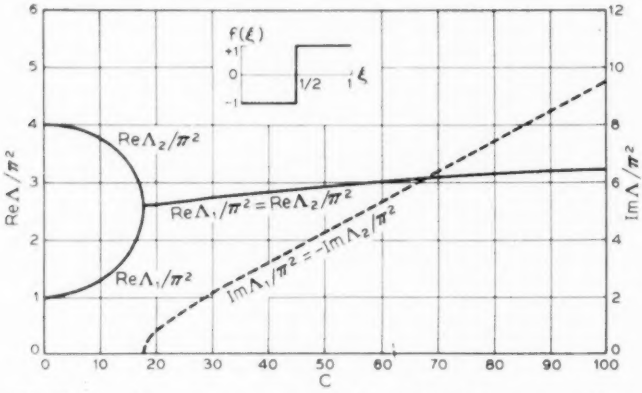


Fig. 22—Real and imaginary parts of Λ_1/π^2 and Λ_2/π^2 for a nonuniform stack whose average properties are constant except at a single, symmetric discontinuity.

(ii) $f(\xi)$ a symmetric rectangular step. Let

$$f(\xi) = \begin{cases} -\dfrac{1}{2\xi_0}, & 0 \leqq \xi < \tfrac{1}{2}\xi_0, \\ \dfrac{1}{2(1-\xi_0)}, & \tfrac{1}{2}\xi_0 < \xi < 1 - \tfrac{1}{2}\xi_0, \\ -\dfrac{1}{2\xi_0}, & 1 - \tfrac{1}{2}\xi_0 < \xi \leqq 1, \end{cases} \tag{546}$$

where ξ_0 is some fixed number between 0 and 1 but not, in the cases of interest, extremely close to either 0 or 1. Inasmuch as $f(\xi)$ has even symmetry about $\xi = \frac{1}{2}$, every mode will preserve the (even or odd) symmetry about $\xi = \frac{1}{2}$ which it has when $C = 0$. We shall consider the lowest even mode,* which has the eigenfunction $\sin \pi\xi$ when $C = 0$. Solutions of (532) having even symmetry about $\xi = \frac{1}{2}$ (we need consider only the region $0 \leqq \xi \leqq \frac{1}{2}$ on account of the symmetry) and satisfying the boundary conditions (533) are given by

* For large C the lowest even mode will be confined essentially to

$$\tfrac{1}{2}\xi_0 < \xi < 1 - \tfrac{1}{2}\xi_0,$$

while the lowest odd mode will be confined to the two regions

$$0 \leqq \xi < \tfrac{1}{2}\xi_0 \text{ and } 1 - \tfrac{1}{2}\xi_0 < \xi \leqq 1.$$

If $\xi_0 > 2/3$, the latter mode will ultimately have a lower attenuation constant than the former; but we shall not take space to investigate it here.

$$\begin{aligned} w(\xi) &= A \sin [\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi, && 0 \leqq \xi < \tfrac{1}{2}\xi_0, \\ w(\xi) &= B \cos [\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}}(\tfrac{1}{2} - \xi), && \tfrac{1}{2}\xi_0 < \xi \leqq \tfrac{1}{2}. \end{aligned} \tag{547}$$

The requirements that w and $dw/d\xi$ must be continuous at $\xi = \frac{1}{2}\xi_0$ lead to the equations

$$\begin{aligned} &A \sin \tfrac{1}{2}[\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0 = B \cos \tfrac{1}{2}[\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}}(1 - \xi_0), \\ &A[\Lambda + iC/2\xi_0]^{\frac{1}{2}} \cos \tfrac{1}{2}[\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0 \\ &\quad = B[\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}} \sin \tfrac{1}{2}[\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}}(1 - \xi_0), \end{aligned} \tag{548}$$

which will be consistent if the following characteristic equation is satisfied:

$$\frac{\tan \tfrac{1}{2}[\Lambda + iC/2\xi_0]^{\frac{1}{2}}\xi_0}{[\Lambda + iC/2\xi_0]^{\frac{1}{2}}} = \frac{\cot \tfrac{1}{2}[\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}}(1 - \xi_0)}{[\Lambda - iC/2(1 - \xi_0)]^{\frac{1}{2}}}. \tag{549}$$

The roots in Λ of equation (549) are the eigenvalues corresponding to the even modes of the symmetrical structure.

When $C = 0$, the roots of (549) are $\Lambda = \pi^2, 9\pi^2, \cdots$. It appears that for $C > 0$ we have $\mathrm{Re}\ \Lambda_1 > \pi^2$ and $\mathrm{Im}\ \Lambda_1 > 0$. For large C the asymptotic expression for Λ_1 turns out to be

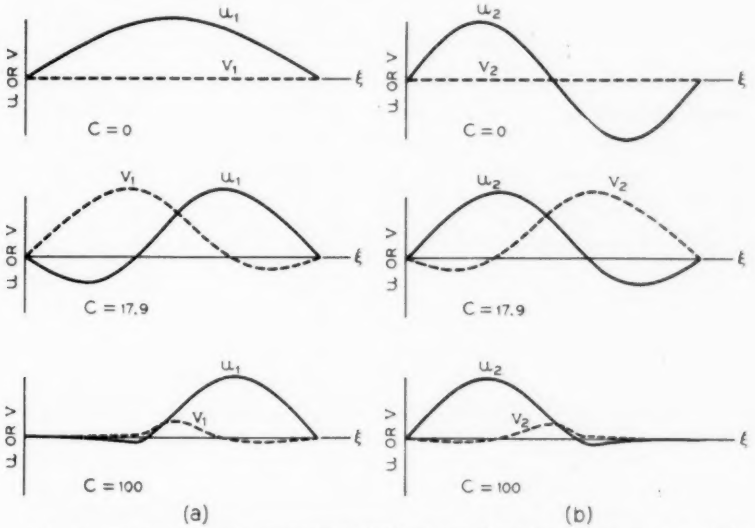


Fig. 23—Real and imaginary parts of the first two eigenfunctions, $w_1 = u_1 + iv_1$ and $w_2 = u_2 + iv_2$, for the nonuniform stack of Fig. 22.

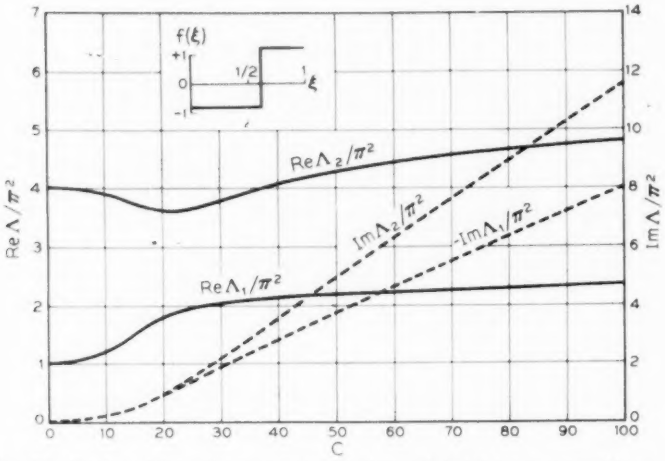


Fig. 24—Real and imaginary parts of Λ_1/π^2 and Λ_2/π^2 for a nonuniform stack whose average properties are constant except at a single, unsymmetric discontinuity.

$$\Lambda \approx \frac{\pi^2}{(1-\xi_0)^2}\left[1 - 4\sqrt{\frac{\xi_0}{(1-\xi_0)C}}\right] + i\left[\frac{C}{2(1-\xi_0)} - \frac{4\pi^2}{(1-\xi_0)^2}\sqrt{\frac{\xi_0}{(1-\xi_0)C}}\right]. \tag{550}$$

Numerical values of Λ_1 were found for the case $\xi_0 = \frac{1}{2}$ on the analog computer and refined afterward by desk computation. The real and imaginary parts of Λ_1/π^2 are plotted over the range $0 \leqq C \leqq 100$ in Fig. 25, and the corresponding eigenfunction $w_1(\xi) = u_1(\xi) + iv_1(\xi)$ is shown in Fig. 26 for $C = 0$, 20, and 100.

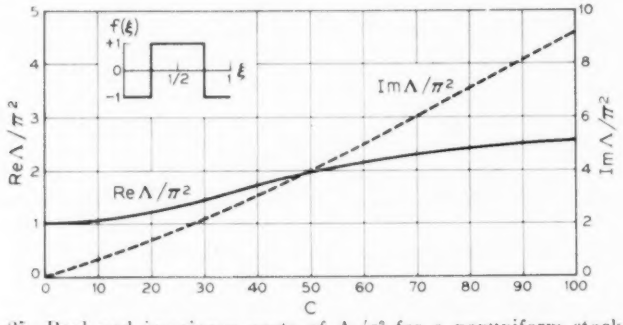


Fig. 25—Real and imaginary parts of Λ_1/π^2 for a nonuniform stack whose average properties vary as a symmetric rectangular step.

(*iii*) $f(\xi)$ *a linear function.* Let

$$f(\xi) = 2\xi - 1, \qquad 0 \leqq \xi \leqq 1, \tag{551}$$

so that $f(\xi)$ is a linear function varying from -1 at $\xi = 0$ to $+1$ at $\xi = 1$. Then equation (532) becomes

$$d^2w/d\xi^2 + [(\Lambda + iC) - 2iC\xi]w(\xi) = 0, \tag{552}$$

which, by the change of variable

$$\tau = \frac{2iC\xi - (\Lambda + iC)}{(2C)^{2/3}}, \tag{553}$$

is transformed into Stokes' equation,

$$d^2w/d\tau^2 + \tau w = 0. \tag{554}$$

The general solution of this equation may be written in the form

$$w = Ah_1(\tau) + Bh_2(\tau), \tag{555}$$

where h_1 and h_2 are the pair of independent solutions of Stokes' equation which have been tabulated for complex arguments by the Computation Laboratory of Harvard University.[27] (The solution may also be ex-

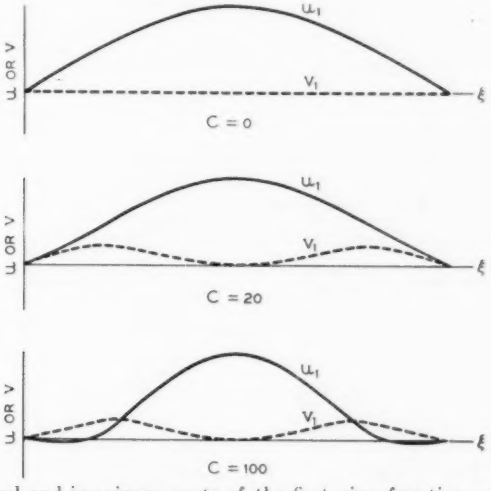


Fig. 26.—Real and imaginary parts of the first eigenfunction, $w_1 = u_1 + iv_1$, for the nonuniform stack of Fig. 25.

[27] *Tables of the Modified Hankel Functions of Order One-Third and of Their Derivatives*, Harvard University Press, Cambridge, Mass., 1945.

pressed, though less conveniently, in terms of Bessel functions of order one-third.) It is easy to show that the boundary conditions (533) at $\xi = 0$ and $\xi = 1$ require

$$\begin{aligned} Ah_1(\tau_1) + Bh_2(\tau_1) &= 0, \\ Ah_1(\tau_2) + Bh_2(\tau_2) &= 0, \end{aligned} \tag{556}$$

where

$$\tau_1 = -\frac{(\Lambda + iC)}{(2C)^{\frac{2}{3}}}, \qquad \tau_2 = -\frac{(\Lambda - iC)}{(2C)^{\frac{2}{3}}}. \tag{557}$$

Equations (556) will be consistent if

$$h_1(\tau_1)h_2(\tau_2) - h_1(\tau_2)h_2(\tau_1) = 0; \tag{558}$$

and this is the relation which must be satisfied by the eigenvalues Λ_1, Λ_2, Λ_3, $\cdots$, for any given value of C.

Approximate values of Λ_1 and Λ_2 have been found using the analog computer for the range $0 \leqq C \leqq 100$, with spot checks by numerical solution of equation (558); and Λ_1/π^2 and Λ_2/π^2 are plotted in Fig. 27. The eigenfunctions are qualitatively similar to those shown in Fig. 23 for the stack with a symmetric discontinuity. As in the symmetric example in case (*i*) above, we find that for small positive C, Λ_1 is real and greater than π^2, while Λ_2 is real and less than $4\pi^2$. The two eigenvalues coincide at

$$C \approx 49, \qquad \Lambda_1 = \Lambda_2 \approx 29. \tag{559}$$

For larger values of C, Λ_1 and Λ_2 are complex conjugates. Their asymp-

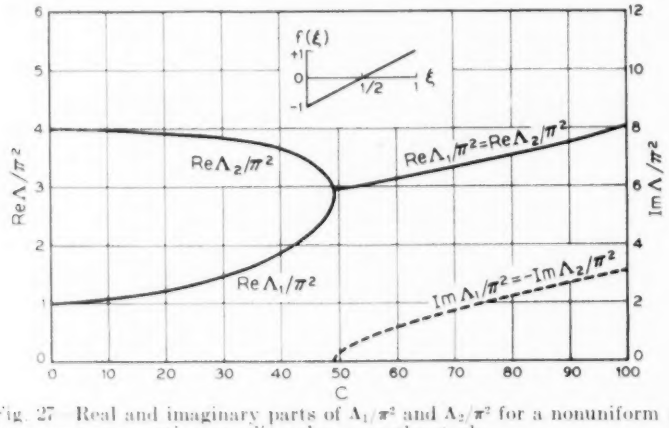


Fig. 27 — Real and imaginary parts of Λ_1/π^2 and Λ_2/π^2 for a nonuniform stack whose average properties vary linearly across the stack.

totic forms as $C \to \infty$ may be deduced by considering the behavior of $h_1(\tau)$ and $h_2(\tau)$ for large arguments, and are

$$\Lambda_1 = \Lambda_2^* \approx 1.169(2C)^{\frac{1}{3}} + i[C - 2.025(2C)^{\frac{1}{3}}]. \tag{560}$$

The magnitudes of both the real and imaginary parts of Λ_1 and Λ_2 thus increase indefinitely with C.

(*iv*) $f(\xi)$ *a sinusoidal function.* Let

$$f(\xi) = -\cos 2\nu\pi\xi, \qquad 0 \leqq \xi \leqq 1, \tag{561}$$

where $\nu = \frac{1}{2}, 1, 2, 3, 4, \cdots$, so that $f(\xi)$ goes through ν complete cycles in $0 \leqq \xi \leqq 1$. Then equation (532) reads

$$d^2w/d\xi^2 + [\Lambda + iC\cos 2\nu\pi\xi]w(\xi) = 0. \tag{562}$$

If we make the transformations

$$\begin{aligned} \tau &= \nu\pi\xi, \\ W(\tau) &= w(\xi), \\ \lambda &= \Lambda/\nu^2\pi^2, \\ \vartheta &= -iC/2\nu^2\pi^2, \end{aligned} \tag{563}$$

we get

$$d^2W/d\tau^2 + [\lambda - 2\vartheta\cos 2\tau]W(\tau) = 0, \tag{564}$$

and the boundary conditions (533) become

$$W(0) = W(\nu\pi) = 0. \tag{565}$$

Equation (564) is one of the standard forms of Mathieu's equation. We are interested in solutions which are periodic with period 2 in ξ, and which approach the form $\sin m\pi\xi$ when $C \to 0$. In terms of τ and ϑ, the function corresponding to the mth mode in the Clogston line must reduce to the form

$$W(\tau) \underset{\vartheta \to 0}{\longrightarrow} \sin \frac{m}{\nu}\tau. \tag{566}$$

For any value of ϑ, this function may be denoted by[28]

$$W(\tau) = \mathrm{se}_{m/\nu}(\tau, \vartheta). \tag{567}$$

[28] See N. W. McLachlan, *Theory and Application of Mathieu Functions*, Oxford, 1947, pp. 10–25, especially p. 13 and p. 19. In this reference a or b corresponds to our λ, q to our ϑ, and ν to our m/ν.

In our problem ϑ is (negative) imaginary and m/ν may be an integer or a rational fraction. For any given ϑ and m/ν the conditions (565) together with the limiting form (566) determine an eigenvalue λ, and hence by (563) determine Λ; but only a small amount of work has been published on the eigenvalues of Mathieu functions with imaginary parameter or of fractional order. We shall look at some special cases.

$\nu = \frac{1}{2}$. The function $f(\xi)$ is one-half cycle of a cosine curve which varies from -1 to $+1$; we expect results similar to those found for the symmetric discontinuity of case (i) and the linear variation of case (iii). The eigenfunctions of the first two modes ($m = 1$ and $m = 2$) are $se_2(\tau, \vartheta)$ and $se_4(\tau, \vartheta)$. The eigenvalues of these two functions for purely imaginary ϑ have been computed by Mulholland and Goldstein[29] out to a point which corresponds to $C = 8\pi^2$, and an asymptotic formula is given for larger values of C. The values of Λ_1/π^2 and Λ_2/π^2 are plotted for $0 \leqq C \leqq 100$ in Fig. 28; the corresponding eigenfunctions resemble those shown in Fig. 23 for the stack with a symmetric discontinuity. Again we find that Λ_1 and Λ_2 are real for small positive C, equal for a particular value of C, and conjugate complex for larger C. The leading terms of the asymptotic formula are, in our notation,

$$\Lambda_1 = \Lambda_2^* \approx [4.7124C^{\frac{1}{2}} - 3.0842 - 1.0901C^{-\frac{1}{2}} - \cdots] + i[C - 4.7124C^{\frac{1}{2}} - 1.0901C^{-\frac{1}{2}} - \cdots]. \tag{568}$$

$\nu = 1$. Here $f(\xi)$ is one full cycle of a cosine function, varying from -1

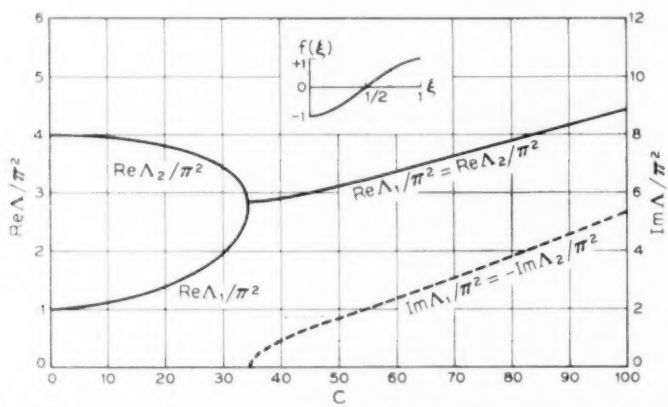


Fig. 28—Real and imaginary parts of Λ_1/π^2 and Λ_2/π^2 for a nonuniform stack whose average properties vary as one-half cycle of a cosine function across the stack.

[29] H. P. Mulholland and S. Goldstein, Phil. Mag. (7), **8**, 834 (1929). In this reference 4α or 4β corresponds to our λ and $8q$ to our ϑ.

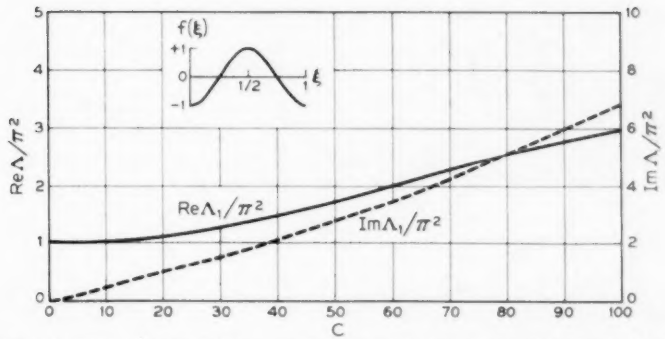


Fig. 29—Real and imaginary parts of Λ_1/π^2 for a nonuniform stack whose average properties vary as one cycle of a cosine function across the stack.

to +1 and back to −1. The eigenfunction of the lowest mode ($m = 1$) is $se_1(\tau, \vartheta)$, and the values of Λ_1 may be obtained from Reference 29 for ten equally spaced values of C out to $C = 32\pi^2$. Since our ϑ is negative imaginary, in the notation of this reference we have $\Lambda_1 = 4\pi^2\beta_1^*$. Approximate values of Λ_1/π^2 obtained on the analog computer for C at smaller intervals in the range $0 \leqq C \leqq 100$ are plotted in Fig. 29; and the eigenfunctions are similar to those shown in Fig. 26 for the symmetric rectangular step. The leading terms of the asymptotic formula for Λ_1 when C is large are as follows:

$$\Lambda_1 \approx [3.1416C^{\frac{1}{2}} - 2.4674 - 0.9689C^{-\frac{1}{2}} - \cdots] + i[C - 3.1416C^{\frac{1}{2}} - 0.9689C^{-\frac{1}{2}} - \cdots]. \tag{569}$$

$\nu = 3$. Now $f(\xi)$ is a three-cycle cosine function and the lowest mode corresponds to $se_1(\tau, \vartheta)$. Approximate values of Λ_1/π^2 for $0 \leqq C \leqq 100$ were obtained on the analog computer and are plotted in Fig. 30; the eigenfunctions are shown in Fig. 31 for $C = 0$ and $C = 100$.

$\nu \gg 1$. For a ν-cycle cosine variation, the lowest eigenfunction is $se_{1/\nu}(\tau, \vartheta)$, and for the lowest eigenvalue there is an approximate formula given by McLachlan.[30] Incidentally this formula predicts no imaginary part for Λ_1 if ϑ is purely imaginary and $\nu > 1$, which agrees approximately with the results of our analog computations for $\nu = 3$; we found the imaginary part of Λ_1 to be only about 1 per cent of the real part even for $C = 100$. If C is fixed, one expects that as $\nu \to \infty$ the effects of the rapid fluctuations in $f(\xi)$ will average out, so that Λ_1 will ultimately approach

[30] Reference 28, p. 20, equation (6), where a corresponds to our λ_1, q to our ϑ, and ν to our $1/\nu$. McLachlan's formula was ostensibly derived for real q, but the derivation appears equally valid for complex q.

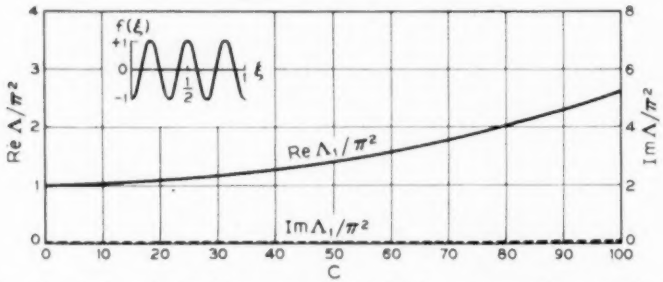


Fig. 30—Real and imaginary parts of Λ_1/π^2 for a nonuniform stack whose average properties vary as three cycles of a cosine function across the stack.

the value π^2 appropriate to a uniform stack. McLachlan's formula shows that this is indeed the case; in our notation, the leading terms give

$$\Lambda \approx \pi^2 + \frac{C^2}{8\nu^2\pi^2} = \pi^2\left[1 + \frac{C^2}{8\nu^2\pi^4}\right], \tag{570}$$

assuming of course that the second term is reasonably small compared to the first.

This concludes our discussion of special types of nonuniformity. We shall now attempt to get an idea of what the numerical results mean in terms of the practical requirements on stack uniformity in a laminated transmission line which is expected to show a specified reduction in attenuation constant below a conventional line of the same dimensions. For this purpose we shall compare a plane Clogston 2 line having infinitesimally thin layers with a plane air-filled line of the same width a, bounded by electrically thick solid conductors.

At frequencies for which the conductor thickness of the "standard"

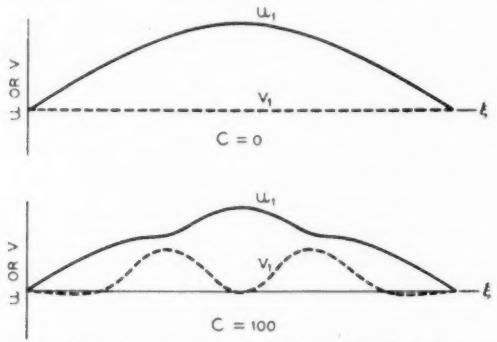


Fig. 31—Real and imaginary parts of the first eigenfunction, $w_1 = u_1 + iv_1$, for the nonuniform stack of Fig. 30.

air-filled line is great compared to the skin depth δ_1, its attenuation constant α_s is given by equation (25), namely

$$\alpha_s = 1/\eta_v g_1 \delta_1 a, \tag{571}$$

where η_v is the intrinsic impedance of free space. By equation (535), the attenuation constant α_c of the lowest mode in a plane Clogston 2 with infinitesimally thin layers is

$$\alpha_c = \text{Re}\, \frac{\Lambda_1}{2\sqrt{\bar{\mu}_0/\bar{\epsilon}_0}\, \bar{g}_0 a^2}. \tag{572}$$

If we assume nonmagnetic materials and put in the optimum value of θ, namely $\theta = 2/3$, we obtain for a uniform stack with $\Lambda_1 = \pi^2$,

$$\alpha_{c0} = \frac{12.82\sqrt{\epsilon_{2r}}}{\eta_v g_1 a^2}, \tag{573}$$

where ϵ_{2r} is the relative dielectric constant of the insulating layers.

The attenuation constant of the conventional line is proportional to the square root of frequency, whereas the attenuation constant of the uniform Clogston 2 is independent of frequency up to some frequency at which the effect of finite lamina thickness begins to be appreciable. If we confine ourselves to the low-frequency, flat attenuation region, and denote the ratio of attenuation constants by r, then from (571) and (573),

$$r = a_{c0}/a_s = 12.82\sqrt{\epsilon_{2r}}\, \delta_1/a, \tag{574}$$

and the crossover frequency above which the uniform Clogston line is better than the conventional line occurs when

$$a/\delta_1 = 12.82\sqrt{\epsilon_{2r}}. \tag{575}$$

In the following numerical example we shall assume polyethylene insulating layers, with

$$\epsilon_{2r} = 2.26, \tag{576}$$

so that (574) becomes

$$r = 19.27\delta_1/a. \tag{577}$$

If the stack in a Clogston line is not uniform, then regardless of the thinness of the layers the attenuation constant will no longer be independent of frequency, but will increase with frequency at a rate depending on the nature and the magnitude of the nonuniformity. Since from equation (572) the attenuation constant is proportional to Re Λ_1,

while from equation (529) or (530), C is proportional to frequency for a given stack, we see that our plots of Re Λ_1/π^2 versus C need only the introduction of appropriate scale factors to read directly the variation of attenuation with frequency due to nonuniformity in the stack. Although a nonuniform Clogston line may still be better under some conditions than a conventional line of the same size, the crossover frequency will be higher and the improvement at any given frequency will be less than if the stack were uniform.

Among the various interpretations which may be given to our numerical results, we shall consider here only the following: Suppose we have a plane Clogston 2 line which, if it were perfectly uniform, would have an attenuation constant smaller, at a certain frequency, than the attenuation constant of the corresponding conventional line by a given factor, say one-half, one-fifth, or one-tenth. For these particular attenuation reduction factors the ratio of a to δ_1 may be calculated from (574), or from (577) if the insulation is polyethylene. The question is: What variation in $\bar{\epsilon}$ across the stack is permissible if we are willing to have the actual attenuation constant of the Clogston line be double its ideal value; in other words, if we will settle for attenuation reduction factors of unity (no improvement), two-fifths, or one-fifth instead of the ideal values one-half, one-fifth, or one-tenth?

To answer this question for any particular type of nonuniformity, we have only to find, from the plot of Re Λ_1/π^2 versus C, the value of C for which Re $\Lambda_1/\pi^2 = 2$. Then the fractional difference between the maximum and minimum values of $\bar{\epsilon}$ corresponding to this value of C is given by equations (530) and (531) to be

$$\frac{\bar{\epsilon}_{\max} - \bar{\epsilon}_{\min}}{\bar{\epsilon}_0} = \frac{3C\delta_1^2}{4a^2}\,(f_{\max} - f_{\min}), \tag{578}$$

where we have taken $\theta_0 = 2/3$, and $f_{\max}$ and $f_{\min}$ are the extreme values of the function $f(\xi)$ which describes the type of nonuniformity being considered.

The special types of nonuniformity which have been studied above fall roughly into three different classes. In four of the cases, namely the symmetric and unsymmetric single discontinuities, the linear variation, and the half-cycle cosine variation, the function $f(\xi)$ varies monotonically from one side of the stack to the other. In the symmetric rectangular step and the one-cycle cosine variation, $f(\xi)$ oscillates from one extreme value to the other and back again, while in the three-cycle cosine variation, $f(\xi)$ exhibits three complete oscillations across the stack. The following table

shows the permissible total variation in $\bar{\epsilon}$ for each of these types of nonuniformity.

Case	C	$\frac{\bar{\epsilon}_{max} - \bar{\epsilon}_{min}}{\bar{\epsilon}_0}$		
		$r = \frac{1}{2}$	$r = \frac{1}{5}$	$r = \frac{1}{10}$
Symmetric discontinuity	16.5	0.0167	0.0027	0.0007
Unsymmetric discontinuity	28.0	0.0295	0.0047	0.0012
Linear	42.6	0.0430	0.0069	0.0017
Half-cycle cosine	29.5	0.0298	0.0048	0.0012
Rectangular step	53.0	0.0535	0.0086	0.0021
One-cycle cosine	59.8	0.0604	0.0097	0.0024
Three-cycle cosine	78.9	0.0797	0.0128	0.0032

It would be easy to construct a similar table for any other values of the attenuation ratio r, and for any specified degradation due to nonuniformity. It is, however, already obvious that the greater the improvement for which one strives, that is, the smaller the ratio r, the more stringent will be the requirement on $(\bar{\epsilon}_{max} - \bar{\epsilon}_{min})/\bar{\epsilon}_0$; in fact, the permissible value of this quantity is proportional to r^2. In any practical case the value of $\bar{\epsilon}$ will have to be controlled against long-range variations within a fraction of a per cent, and if attenuation reduction factors of the order of one-fifth or one-tenth are contemplated, the variations probably cannot exceed a few hundredths of a per cent. It also appears that a steady increase or decrease in the value of $\bar{\epsilon}$ across the stack will be the most serious type of nonuniformity, since the effects of very rapid fluctuations will tend to average out.

Clearly the nonuniform laminated transmission lines which we have been considering in this section are very highly idealized, even if we disregard the geometrical differences between plane and coaxial structures. Any real Clogston cable will be built up of layers of finite thickness with unavoidable random fluctuations from layer to layer, superimposed on slower variations in the average properties of the layers from one side of the stack to the other. The thickness of an individual layer will also vary more or less in both directions parallel to the layer, so that the properties of the stack will be functions of the coordinates ϕ and z as well as of ρ. A few qualitative remarks are in order concerning these neglected effects.

The effect of finite lamina thickness in a nonuniform stack can be calculated, by the method employed in Section XI for a uniform coaxial stack, if we make the plausible assumption that the macroscopic current distribution remains the same as for infinitesimally thin layers. The results

will certainly be qualitatively the same for uniform and slightly nonuniform stacks, so long as the nonuniformity does not seriously distort the field pattern of the operating mode.

Some idea of how the effects of rapid random fluctuations in the average properties of the stack may be expected to average out is given by equation (570), which assumes for the function $f(\xi)$ a ν-cycle cosine variation across the stack. As a numerical example, suppose that with this variation of $\bar{\epsilon}$ we have

$$\frac{\bar{\epsilon}_{\max} - \bar{\epsilon}_{\min}}{\bar{\epsilon}_0} = 0.01 \tag{579}$$

in a line designed to give an attenuation reduction ratio of

$$r = \tfrac{1}{10}. \tag{580}$$

Assuming polyethylene insulating layers, we have for this line

$$\delta_1/a = 1/192.7 = 0.00519, \tag{581}$$

and from (578) the corresponding value of C is

$$C = 247.6. \tag{582}$$

The value of ν for which the relative increase in Re Λ_1 due to the fluctuations is, say, one-quarter is given by

$$\frac{C^2}{8\nu^2\pi^4} = \frac{1}{4}, \qquad \nu = \frac{C}{\sqrt{2}\,\pi^2} = 17.7. \tag{583}$$

Thus a 1 per cent fluctuation in $\bar{\epsilon}$, repeated at intervals of about one-eighteenth of the stack width, will cause only a 25 per cent increase in attenuation, even for a Clogston line which is designed to have only one-tenth of the attenuation constant of a conventional line of the same size.

Finally there is the question of the effects of variations in the average properties of the stack in both directions parallel to the layers. Mathematical analysis of even a simple case of longitudinal variation would be much more difficult than what has been done here; yet on physical grounds it seems very likely that such variations will add an appreciable amount to the total attenuation of the line. If we consider two cross sections of a laminated cable separated by a certain distance and having different transverse nonuniformities, the field pattern of the lowest mode will be different at the two cross sections, and so in traversing the intervening distance the power will be partly reflected and partly converted to higher modes with higher attenuation constants. The reflected or mode converted power will be at least partly lost, with a consequent increase in the overall at-

tenuation of the cable. Hence the estimate of the increase in attenuation which one gets from the present analysis, considering only the variations transverse to the layers at an average cross section, is certain to be optimistic in that it neglects completely the effects of variations in other directions.

XIII. DIELECTRIC AND MAGNETIC LOSSES IN CLOGSTON 2 LINES

To discuss dielectric and magnetic losses in Clogston 2 lines we may take the electrical constants of the conducting and insulating layers to be complex; thus

$$\begin{aligned} \mu_1 &= \mu_1' - i\mu_1'' = \mu_1'(1 - i \tan \zeta_1), \\ \mu_2 &= \mu_2' - i\mu_2'' = \mu_2'(1 - i \tan \zeta_2), \\ \epsilon_2 &= \epsilon_2' - i\epsilon_2'' = \epsilon_2'(1 - i \tan \phi_2). \end{aligned} \tag{584}$$

Almost all of the equations of the preceding sections, except of course those which involve explicit separation of real and imaginary parts, remain valid when we introduce complex values of μ_1, μ_2, and ϵ_2. In particular the propagation constant of the pth mode in a Clogston 2 with infinitesimally thin laminae and high-impedance walls is given, as in Sections VIII through X, by

$$\gamma^2 = -\omega^2\bar{\mu}\bar{\epsilon} + (i\omega\bar{\epsilon}/\bar{g})\chi_p^2, \tag{585}$$

where

$$\chi_p = p\pi/a \tag{586}$$

for a parallel-plane line, and χ_p is the pth root of

$$J_1(\chi a)N_1(\chi b) - J_1(\chi b)N_1(\chi a) = 0 \tag{587}$$

for a coaxial line. Taking the square root of the right side of (585) by the binomial theorem, we have

$$\gamma = i\omega\sqrt{\bar{\mu}\bar{\epsilon}} + \frac{\chi_p^2}{2\sqrt{\bar{\mu}/\bar{\epsilon}}\,\bar{g}}. \tag{588}$$

In the presence of dielectric and/or magnetic dissipation, we write, as in Section VII,

$$\begin{aligned} \bar{\epsilon} &= \bar{\epsilon}' - i\bar{\epsilon}'' = [\epsilon_2'/(1 - \theta)] - i[\epsilon_2''/(1 - \theta)], \\ \bar{\mu} &= \bar{\mu}' - i\bar{\mu}'' = [\theta\mu_1' + (1 - \theta)\mu_2'] - i[\theta\mu_1'' + (1 - \theta)\mu_2'']. \end{aligned} \tag{589}$$

Then the term $i\omega\sqrt{\bar{\mu}\bar{\epsilon}}$ in (588) has a small real part, namely

$$\alpha_d = \tfrac{1}{2}\omega\sqrt{\bar{\mu}'\bar{\epsilon}'}\,(\tan\phi_2 + \tan\bar{\xi}), \tag{590}$$

where

$$\tan\bar{\xi} = \frac{\bar{\mu}''}{\bar{\mu}'} = \frac{\theta\mu_1'' + (1-\theta)\mu_2''}{\theta\mu_1' + (1-\theta)\mu_2'}; \tag{591}$$

and α_d is the part of the attenuation constant which is due to dielectric and magnetic losses. If there were no dielectric or magnetic dissipation, the second term on the right side of (588) would be purely real and would represent the attenuation due to ohmic losses in the conducting layers. We neglect the small change in this term when $\bar{\mu}$ and $\bar{\epsilon}$ are complex, and thus as usual regard the metal losses, the dielectric losses, and the magnetic losses as additive.

We observe that α_d is the same for both plane and coaxial lines, and is also independent of the mode number p. Although derived here for the case of infinitesimally thin laminae, the same expression may be used for lines with finite laminae, so long as the conducting layers are moderately thin compared to the skin depth. The dielectric and magnetic losses do not depend on the overall dimensions of the transmission line, but are directly proportional to frequency provided that the loss tangents do not vary with frequency.

If it should be necessary to calculate the dielectric and magnetic losses in a partially filled Clogston line where the dissipation factor of the main dielectric is markedly different from the dissipation factor of the stacks, α_d may be obtained, using the method described in Section VII, as half the ratio of dissipated power per unit length to transmitted power. In this calculation we may use the field components given in Sections IX and X for the various modes in partially filled lines.

ACKNOWLEDGMENTS

Many people with whom I have discussed the theoretical and practical aspects of the laminated transmission line problem at various times have offered comments and suggestions which are reflected in this paper. I have especially to express my appreciation to A. M. Clogston, H. S. Black, and J. G. Kreer, Jr., for stimulating and helpful discussions.

My thanks are also extended to Mrs. M. F. Shearer, Mrs. D. R. Fursdon, and Miss R. A. Weiss for the extensive numerical computations which they carried out in connection with this study, and to Miss D. T. Angell for preparation of the curves and diagrams.

Appendix II

Optimum proportions for heavily loaded Clogston cables

We wish to find the lowest root χ_1 of the equation

$$\frac{1}{\chi\rho_1}\frac{J_1(\chi a)N_0(\chi\rho_1) - N_1(\chi a)J_0(\chi\rho_1)}{J_1(\chi a)N_1(\chi\rho_1) - N_1(\chi a)J_1(\chi\rho_1)} + \frac{1}{\chi\rho_2}\frac{J_1(\chi b)N_0(\chi\rho_2) - N_1(\chi b)J_0(\chi\rho_2)}{J_1(\chi\rho_2)N_1(\chi b) - N_1(\chi\rho_2)J_1(\chi b)} = \frac{\mu_0}{\bar{\mu}}\log\frac{\rho_2}{\rho_1}, \tag{A9}$$

where b is fixed and $\mu_0/\bar{\mu} \gg 1$, and to minimize this root as a function of a, ρ_1, and ρ_2.

Since we expect χ_1 to approach zero as $\mu_0/\bar{\mu}$ approaches infinity, we shall replace the Bessel functions appearing in (A9) by their approximate values for small argument, namely

$$\begin{aligned} J_0(x) &\approx 1, \\ J_1(x) &\approx \tfrac{1}{2}x, \\ N_0(x) &\approx \frac{2}{\pi}\log 0.8905x, \\ N_1(x) &\approx -\frac{2}{\pi x}, \end{aligned} \tag{A10}$$

for $|x| \ll 1$. Then the equation becomes, approximately,

$$\frac{1}{\chi_1\rho_1}\frac{1/\chi_1 a}{\tfrac{1}{2}(-a/\rho_1 + \rho_1/a)} + \frac{1}{\chi_1\rho_2}\frac{1/\chi_1 b}{\tfrac{1}{2}(-\rho_2/b + b/\rho_2)} = \frac{\mu_0}{\bar{\mu}}\log\frac{\rho_2}{\rho_1}, \tag{A11}$$

which may be solved for χ_1^2 to yield

$$\chi_1^2 = \frac{2\bar{\mu}}{\mu_0 \log(\rho_2/\rho_1)}\left[\frac{1}{\rho_1^2 - a^2} + \frac{1}{b^2 - \rho_2^2}\right]. \tag{A12}$$

By inspection χ_1^2 will be a minimum, considered as a function of a, when $a = 0$. Setting $a = 0$ and then equating to zero the partial derivatives of χ_1^2 with respect to ρ_1 and ρ_2, we get the pair of equations

$$\begin{aligned} \frac{1}{\rho_1[\log(\rho_2/\rho_1)]^2}\left[\frac{1}{\rho_1^2} + \frac{1}{b^2 - \rho_2^2}\right] - \frac{2}{\rho_1^3 \log(\rho_2/\rho_1)} &= 0, \\ -\frac{1}{\rho_2[\log(\rho_2/\rho_1)]^2}\left[\frac{1}{\rho_1^2} + \frac{1}{b^2 - \rho_2^2}\right] + \frac{2\rho_2}{(b^2 - \rho_2^2)^2 \log(\rho_2/\rho_1)} &= 0, \end{aligned} \tag{A13}$$

which yield, on rearrangement,

$$\frac{1}{\log (\rho_2/\rho_1)}\left[\frac{1}{\rho_1^2}+\frac{1}{b^2-\rho_2^2}\right]=\frac{2}{\rho_1^2},$$
$$\frac{1}{\log (\rho_2/\rho_1)}\left[\frac{1}{\rho_1^2}+\frac{1}{b^2-\rho_2^2}\right]=\frac{2\rho_2^2}{(b^2-\rho_2^2)^2}. \tag{A14}$$

Subtracting the second of equations (A14) from the first and solving the resulting equation for ρ_1, we get

$$\rho_1=\frac{b^2-\rho_2^2}{\rho_2}=\frac{b^2}{\rho_2}-\rho_2, \tag{A15}$$

whence, eliminating ρ_1 from the first of (A14),

$$\log\frac{\rho_2^2}{b^2-\rho_2^2}=\frac{b^2}{2\rho_2^2}. \tag{A16}$$

Numerical solution of (A16) gives

$$\rho_2^2/b^2=0.67674; \tag{A17}$$

and on making use of (A15) we obtain finally

$$\rho_1=0.39296b, \tag{A18}$$
$$\rho_2=0.82264b.$$

Substituting these values, with $a=0$, into (A12), we get for the minimum value of χ_1^2, when $\mu_0/\bar{\mu}\gg 1$,

$$\chi_1^2=\frac{25.905\bar{\mu}}{\mu_0 b^2}. \tag{A19}$$

Appendix III

Power Dissipation in a Hollow Conducting Cylinder

Consider a hollow cylinder of inner radius ρ_1, outer radius ρ_2, and high conductivity g_1. Denote the total current flowing in the positive z-direction inside the radius ρ_1 by I_1 and the total current inside the radius ρ_2 by I_2; then the current carried by the conducting cylinder is just I_2-I_1, and the net return current outside the cylinder is $-I_2$. We assume the current distribution to be independent of the coordinate angle ϕ, but the radial distribution of the currents inside and outside the given cylinder is of no importance.

General expressions for the field components in the conducting cylinder

are given by equations (33) of Section II, which read

$$\begin{aligned} H_\phi &= AI_1(\sigma_1\rho) + BK_1(\sigma_1\rho), \\ E_\rho &= \frac{\gamma}{g_1}[AI_1(\sigma_1\rho) + BK_1(\sigma_1\rho)], \\ E_z &= \eta_1[AI_0(\sigma_1\rho) - BK_0(\sigma_1\rho)], \end{aligned} \tag{A20}$$

provided that we drop the propagation factor $e^{-\gamma z}$ and make the usual approximations

$$g_1/\omega\epsilon_1 \gg 1, \qquad \kappa_1 \approx \sigma_1, \qquad \kappa_1/g_1 \approx \eta_1, \tag{A21}$$

for a good conductor. The constants A and B are determined by the boundary conditions

$$H_\phi(\rho_1) = I_1/2\pi\rho_1, \qquad H_\phi(\rho_2) = I_2/2\pi\rho_2, \tag{A22}$$

which follow directly from Ampere's circuital law. We find without difficulty

$$\begin{aligned} A &= \frac{(I_2/2\pi\rho_2)K_1(\sigma_1\rho_1) - (I_1/2\pi\rho_1)K_1(\sigma_1\rho_2)}{K_1(\sigma_1\rho_1)I_1(\sigma_1\rho_2) - K_1(\sigma_1\rho_2)I_1(\sigma_1\rho_1)}, \\ B &= \frac{(I_1/2\pi\rho_1)I_1(\sigma_1\rho_2) - (I_2/2\pi\rho_2)I_1(\sigma_1\rho_1)}{K_1(\sigma_1\rho_1)I_1(\sigma_1\rho_2) - K_1(\sigma_1\rho_2)I_1(\sigma_1\rho_1)}. \end{aligned} \tag{A23}$$

The average power dissipated in the conducting cylinder is equal to one-half the real part of the inward normal flux of the complex Poynting vector $\mathbf{E} \times \mathbf{H}^*$. For the average power P dissipated per unit length we have

$$\begin{aligned} P &= \text{Re } \tfrac{1}{2}[2\pi\rho_2 E_z(\rho_2)H_\phi^*(\rho_2) - 2\pi\rho_1 E_z(\rho_1)H_\phi^*(\rho_1)] \\ &= \text{Re } \tfrac{1}{2}[E_z(\rho_2)I_2^* - E_z(\rho_1)I_1^*] \\ &= \text{Re } \frac{\tfrac{1}{2}\eta_1}{(K_{11}I_{12} - K_{12}I_{11})}\left[\frac{I_2I_2^*}{2\pi\rho_2}(K_{11}I_{02} + K_{02}I_{11})\right. \\ &\qquad \left. - \frac{(I_1I_2^* + I_1^*I_2)}{2\pi\sigma_1\rho_1\rho_2} + \frac{I_1I_1^*}{2\pi\rho_2}(K_{01}I_{12} + K_{12}I_{01})\right], \end{aligned} \tag{A24}$$

where

$$I_{rs} = I_r(\sigma_1\rho_s), \qquad K_{rs} = K_r(\sigma_1\rho_s). \tag{A25}$$

The combinations of Bessel functions appearing in (A24) are just those for which we gave approximate expressions in equations (A8) of Appendix I, assuming the thickness t_1 $(= \rho_2 - \rho_1)$ of the conducting cylinder to be small compared to ρ_1. Substituting these approximations into (A24) and

rearranging, we get

$$P = \mathrm{Re}\,\frac{1}{4\pi g_1 t_1}\left\{\frac{I_2 I_2^*}{\rho_2}\left[\sigma_1 t_1 \coth \sigma_1 t_1 + \frac{t_1}{2\rho_1}\right]\right.$$
$$- \frac{(I_1 I_2^* + I_1^* I_2)}{\rho_1}\left[1 - \frac{t_1}{2\rho_1}\right]\sigma_1 t_1 \operatorname{csch} \sigma_1 t_1 \qquad \text{(A26)}$$
$$\left. + \frac{I_1 I_1^*}{\rho_1}\left[\sigma_1 t_1 \coth \sigma_1 t_1 - \frac{t_1}{2\rho_1}\right]\right\},$$

up to first order in t_1/ρ_1 . If we set

$$\sigma_1 = (1 + i)/\delta_1 , \qquad \text{(A27)}$$

then on expanding the right side of (A26) in powers of t_1/δ_1 up to the fourth, we obtain

$$P = \frac{1}{4\pi g_1 t_1}\left\{\frac{|\, I_2 - I_1\,|^2}{\sqrt{\rho_1\rho_2}}\right.$$
$$\left. + \frac{4t_1^4}{\delta_1^4}\left[\frac{I_2 I_2^*}{45\rho_2} + \frac{7(I_1 I_2^* + I_1^* I_2)}{360\sqrt{\rho_1\rho_2}} + \frac{I_1 I_1^*}{45\rho_1}\right]\right\}, \qquad \text{(A28)}$$

where we have approximated $\rho_1(1 + t_1/2\rho_1)$ by $\sqrt{\rho_1\rho_2}$ in the interest of symmetry.

Now writing ΔP_j for P, ΔI_j for $I_2 - I_1$, ρ_{j-1} for ρ_1 , and I_{j-1} for I_1 , and neglecting curvature corrections of the order of t_1/ρ_1 entirely, we have, on setting I_1 and I_2 both equal to I_{j-1} in the coefficient of $(t_1/\delta_1)^4$, the approximate relation

$$\Delta P_j = \frac{1}{4\pi g_1 t_1 \rho_{j-1}}\left[|\,\Delta I_j\,|^2 + \frac{t_1^4}{3\delta_1^4}\,|\,I_{j-1}\,|^2\right], \qquad \text{(A29)}$$

which is just equation (472) of Section XI.

Transistors in Switching Circuits

By A. EUGENE ANDERSON

(Manuscript received August 1, 1952)

The general transistor properties of small size and weight, low power and voltage, and potential long life suggest extensive application of transistors to pulse or switching type systems of computer or computer-like nature. It is possible to devise simple regenerative circuits which perform the normally employed functions of waveform generation, level restoration, delay, storage (registry or memory), and counting. The discussion is limited to point contact type transistors in which the alpha or current gain is in excess of unity and to a particular feedback configuration. Such circuits, which are of the so-called trigger type, are postulated to involve negative resistance. On this basis an analysis, which approximates the negative resistance characteristic by three intersecting broken lines, is developed. Conclusions which are useful to circuit and device design are reached. The analysis is deemed sufficiently accurate for the first order equilibrium calculations. Transistors having properties specifically intended for pulse service in the circuits described have been developed. Their properties, and limitations, and parameter characterizations are discussed at some length.

INTRODUCTION

It is proposed to discuss some of the properties of transistors which are applicable to switching or pulse-type circuits, to develop elementary analysis methods and to describe a few circuits.

The bounds or limits of the field of switching are difficult to define. The common thread usually involves definite states of being as "open or closed", "off or on", "0 or 1", and so on, rather than a continuum of conditions. Even when consideration is given to more than two states, the thought involves distinct recognition of each state. The field is termed to be non-linear in distinction to linear manipulation of information. Any number of anomalies in definition may be raised.

Without attempting either to define or to limit the field, some of the functions which are often employed are: wave form generation, as rectangular pulses, sawtooth waves, etc.; memory or storage which may

be for short, intermediate or long periods and involves the retention of information for subsequent use; operations involving addition, subtraction, multiplication and division; translation of information from one form or code to another; gating, involving the routing of signals according to a predetermined pattern or set of conditions; regeneration of signals in amplitude and wave form; delay, which may be thought of as a form of storage; and timing. Some of these functions are simple; others result from fairly complex structures of simpler functions.

Present trends in electronic switching systems are toward complicated automata as exemplified by digital computers.[1] The reliability, power consumption and physical size of the electron devices employed largely determines the degree of realizability of such systems. It is believed that the transistor will find a significant application in this field.

The transistor can reduce power consumption by the elimination of heater or filament power. In addition, particularly in broadband applications as in high speed pulse systems, the "B" power may be reduced by the order of one or two decades if not more. Transistor circuits with 0.02 μs rise time have been made to operate with an input power of 20 milliwatts which compares with approximately 2.5 watts (1-watt heater, 1.5-watt plate) for an equivalent tube circuit. Transistors have operated with less than one microwatt input power.[2]

Such power reductions result from the low operating voltages, low internal resistances and low capacitances of transistors. Low internal impedances greatly reduce the importance of stray wiring capacitances thereby making mechanical design much simpler and often eliminating the need for isolating or buffer amplifiers.

The transistor can contribute definite reduction in size directly. Fig. 1 shows a "bead" transistor which has a volume of approximatelv 1/1000 of a cubic inch and a weight of 5/1000 ounce. Indirectly the transistor can contribute to size reduction through the use of smaller, lower voltage, lower dissipation components. The reduction of power supply requirements in terms of size, regulation and capacity is also quite appreciable.

Transistors have been subjected to shocks in excess of 20,000 G without change in characteristics. Vibration tests have shown no resonances in the transistor shown in Fig. 1 to several thousand cycles. Harmonic accelerations of 100 G at 1000 cycles have produced no detectable current modulation.

[1] L. N. Ridenour, "High Speed Digital Computers", *J. Appl. Phys.*, **21**, pp. 263–270, April, 1950.

[2] R. L. Wallace, Jr. and W. J. Pietenpol, "Some Circuit Properties and Applications of *n-p-n* Transistors", *Bell System Tech. J.*, **30**, pp. 530–563, July, 1951.

Life reliability and expectancy are difficult to determine due to the relative infancy of transistors, the definite finiteness of time, the many variables involved and the rate of development progress. Average life is presently estimated to be in excess of 70,000 hours. Life is a function of the operating conditions and may be materially reduced accordingly.

Transistors also have limitations. Noise at present is high for point-contact types as compared to electron tubes; input impedances are low, which may be either advantageous or disadvantageous; power output may be limited; frequency response is relatively low; circuit instability may cause design difficulties; and the devices are sensitive to temperature changes. There is also an absence of a long practical experience with a consequent art background in both devices and circuits.

A comprehensive review of transistor properties is given in the paper by J. A. Morton.[3]

While it is difficult to define the switching field, it is no less difficult to discuss circuit and device properties on a general basis. This is related to the non-linear nature of the circuits and devices in distinction to the virtually classical linear small-signal field. The lack of a classical method of analysis is a serious handicap in the synthesis of contemporary circuits and devices. When new devices, as the transistor, are to be considered,

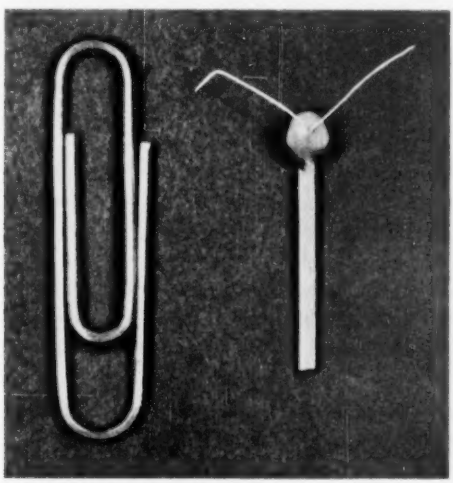


Fig. 1—A miniature switching-type transistor (M1689).

[3] J. A. Morton, "Present Status of Transistor Development", *Bell System Tech. J.*, **31,** pp. 411–442, May, 1952.

the problem is multiplied due to the lack of a long background of experience.

It has been assumed that negative resistance is a common thread among "trigger circuits" and oscillators regardless of the device employed — electron tube, gas tube, transistor, mechanical structures, etc. This is not a new or novel idea and there is no intent to present it as such.[4] Rather, it is used as a pattern upon which a certain degree of transistor analysis may be based, leading to simple understanding. The analysis assumes that the negative resistance characteristic can be broken into three regions; each region is then considered on a linear basis.

Section I will deal with simple circuit properties; Section II with analysis and Section III with device properties.

I—Simple Circuit Properties

The common property ascribable to switching functions is that of definiteness of state. The condition of the function is either "off" or "on". Switches are either open or closed; relays are operated or not; tubes are in cutoff or overload; doors are open or closed and so on. This is common regardless of the phenomena being exploited.

There is an intermediate region between these two conditions usually characterized by a time which is related to how fast the function may go from one state to another. Functionally the times of closing and opening are taken to be zero; practically, they are of determining importance. Relays replace hand-operated switches and electronic devices replace relays as speed becomes important. Obviously, no function or system can be faster than its state-devices.

All such state-devices will have separate attendant properties such as the degree of reverse coupling between the controlling signal and the controlled signal. Separated into families, however, there are those which are passive and those which are active. The latter are threshold devices in which a small amount of signal or control energy causes the translation of a relatively larger amount of stored energy into dynamic energy which consummates the change in state. As long as the control

[4] See for example "Negative Resistances, Their Characteristics and Effects. Sinusoids, Relaxation Oscillations and Relaxation Discontinuities", Walter Reichardt, *Elektrische Nachtrichen-Technik*, **20**, pp. 76–87, March, 1943; "Uniform Relationship Between Sinusoids, Relaxation Vibrations and Discontinuities", Walter Reichardt; *Elektrische Nachtrichten-Technik*, **20**, pp. 213–225, Sept., 1943. For transistors: "Counter Circuits Using Transistors", E. Eberhard, R. O. Endres and R. P. Moore, *RCA Review*, pp. 459–476, Dec. 1949; "A Transistor Trigger Circuit", H. J. Reich and Ungvary, *Rev. Sci. Instr.*, **20**, p. 8, p. 586, Aug., 1949; and "Some Transistor Trigger Circuits", *Proc. Inst. Radio Engrs.*, **39**, pp. 627–632, June, 1951, P. M. Schultheiss and H. J. Reich.

signal is below the initial threshold there is no response and any change is directly related to the passive transmission of the control signal alone. When the signal exceeds the threshold the second state is assumed. Watch escapements, thyratrons, and the whole family of oscillators fall into this category. When the simplest cases of such functions are analyzed, they are found to involve in one way or another two stable states separated by a region in which there is positive feedback and gain in excess of unity with a resultant equivalent negative resistance. The proposition that a negative resistance characteristic is common to trigger or threshold switching circuits is tacitly assumed. The next step is to examine the transistor for such behavior and to classify the properties.

NEGATIVE RESISTANCE IN THE TRANSISTOR

That the transistor* can exhibit negative resistance has been demonstrated analytically[5] and experimentally. The resistances seen looking into the emitter and collector of the transistor with grounded base are shown in Fig. 2.

In the equations and discussion to follow, the symbol conventions are as follows: External circuit elements are capitalized as R_e, R_b, and R_c. The symbols R_{11}, R_{12}, R_{22} and R_{21} define the open-circuit transistor resistances; the symbols r_e, r_c, r_m, and r_b define the equivalent circuit

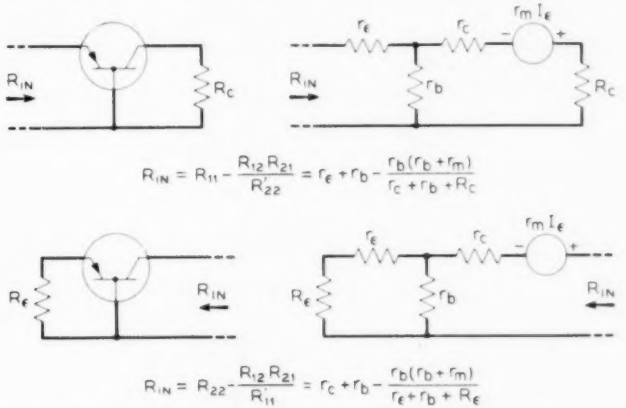


Fig. 2—Emitter and collector driving point resistances.

* Discussion is limited primarily to point contact transistors with α's or current gains greater than unity.

[5] R. M. Ryder and R. J. Kircher, "Some Circuit Aspects of the Transistor", *Bell System Tech. J.*, **28,** pp. 367–400, July, 1949.

element values. Network resistances which contain both device and external elements are primed. For example, $R'_{22} = R_{22} + R_c + R_b$, where $R_{22} = r_c + r_b$. See also references 3 and 5.

Taking the collector or output resistance, Fig. 2, for example,

$$R_{in} = (r_c + r_b) - \frac{r_b(r_b + r_m)}{r_e + r_b + R_e} \tag{1}$$

R_{in} can be negative or positive depending upon the relative magnitudes of the two terms. Actually, of course, r_m has a phase factor and so is frequency dependent. Frequencies wherein r_m is essentially resistive will be assumed. For negative resistance, r_m must be large, R_e small and r_b not too small or else augmented by external resistance. Negative resistance is thus predicted on a small-signal linear basis. The large-signal behavior may be studied experimentally by adding sufficient resistance as R_c to the first or positive term to insure stability. This is shown in Fig. 3 with the resultant characteristic. External base resistance R_b has been added and R_e is zero.

Fig. 3 illustrates the pattern of a three-valued characteristic: Regions I and III are portions with positive slope, indicating stable operating

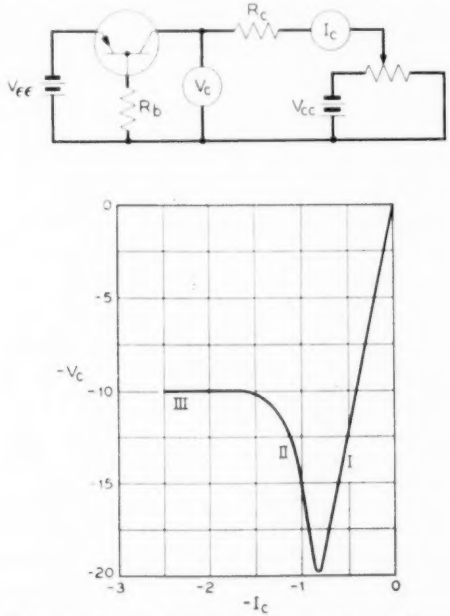


Fig. 3—Collector large-signal negative resistance characteristic.

regions, separated by Region II, a region of negative slope, indicating the possibility of instability. In this particular case, Region I has high resistance and Region III very low resistance.

An evaluation of the emitter or input characteristic leads to similar results, using the circuit of Fig. 4. R_b has been added here also and R_c taken as zero. The general pattern is again present. Region I has high, positive resistance; Region II, negative resistance; and Region III very low, positive resistance.

BIASES AND LOAD LINES—BISTABLE OPERATION

The negative resistances of Figs. 3 and 4 are both of the so-called open-circuit stable type. If loads are applied to the circuit terminals of Fig. 2 which are larger in magnitude than the negative resistances, the circuits will be stable; that is, there will be single operating points. This is shown in Fig. 4 by the dashed load lines marked, R_ϵ', R_ϵ'', R_ϵ'''. A load resistance smaller in magnitude than the negative resistance may intersect the characteristic in three positions as shown by the load line R_ϵ.

The load line R_ϵ can be made to have single or multiple intersections by biasing properly as shown in Fig. 5, where the three possibilities are shown as R_ϵ, R_ϵ', R_ϵ''. Single or multiple intersections result in accordance with the choice of emitter bias, $V_{\epsilon\epsilon}$, as shown. It can be shown

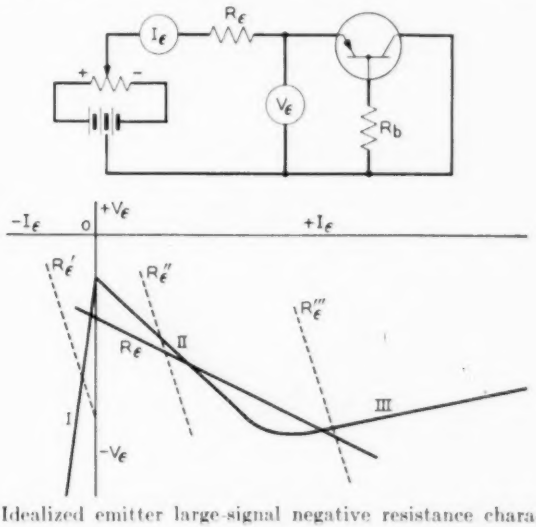


Fig. 4—Idealized emitter large-signal negative resistance characteristic.

that the intersection of load line R_L with the characteristic at b in Fig. 5 is unstable whereas those at a and c are stable. Experiment in the multiple intersection case shows also that as V_{cc} is slowly increased (decreased in absolute magnitude) the load line moves upward and that the assumed operating point, a, moves up along the Region I portion of the characteristic. At the turning point shown on the current axis, the operating point suddenly flips to the high current region, returning along the curve to c as V_{cc} is returned to the original value.

A decrease in V_{cc} toward V_{cc}'' moves the operating point at c downward along the characteristic until it "escapes" past the lower turning point and flips to the Region I portion, returning to a as V_{cc}'' is returned to the original value. This then is an elementary switching circuit, a bistable trigger circuit or "flip-flop". A positive emitter pulse will cause the circuit to flip to high current, a negative pulse to low current. The triggers may be applied to emitter, base or in combination with proper attention to polarity. Trigger sensitivities are shown in Fig. 6. Such a circuit is often used for register or storage purposes. It can store one bit of information for a potentially infinite period, be sampled for the presence of such information, and be cleared or restored to the original condition for reuse when the stored information is no longer useful.

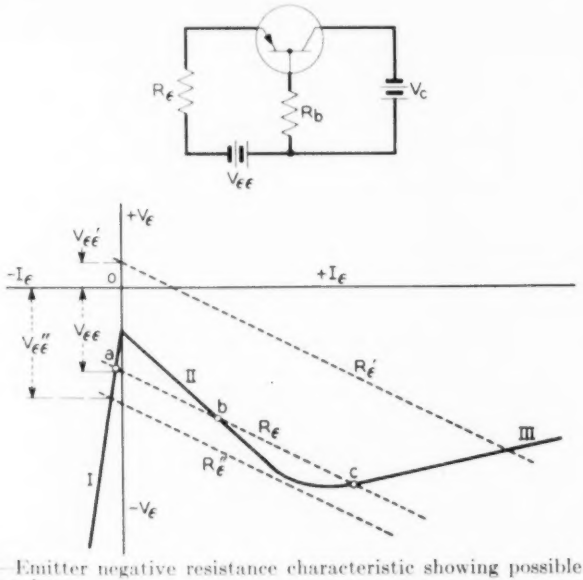


Fig. 5—Emitter negative resistance characteristic showing possible multiple operating points.

With the addition of suitable steering circuits it can be made to count by a scale of two.

MONOSTABLE AND ASTABLE CIRCUITS

The addition of a capacitor to the circuit as in Fig. 7(a) leads to either monostable or astable operation. In Fig. 7(b) the normal operating point is stable at a as discussed previously by virtue of the bias V_{ee}. As V_{ee} is increased, as by a trigger, the load line is moved up and over the turning point. Without capacitor C in the circuit, the operating point would move to b with the resultant rapid change in voltage and current. However, a capacitor has in effect voltage inertia; this is equivalent to saying that a capacitor is a short-circuit to a voltage change. Both the capacitance and the rate of change of voltage are assumed high. Thus at the turning point the capacitor effectively short-circuits the emitter and the operating point snaps along dotted line (1) to intersect the characteristic. This point is quasi-stable and the capacitor is discharged along line (2) to the second turning point where the emitter is again effectively short-circuited and the operating point snaps along (3) to intersect the Region I portion of the characteristic. This point is also quasi-stable and the operating point moves slowly up to the initial or dc stable operating point. A single trigger thus causes a complete cycle of operation. The emitter current shifts

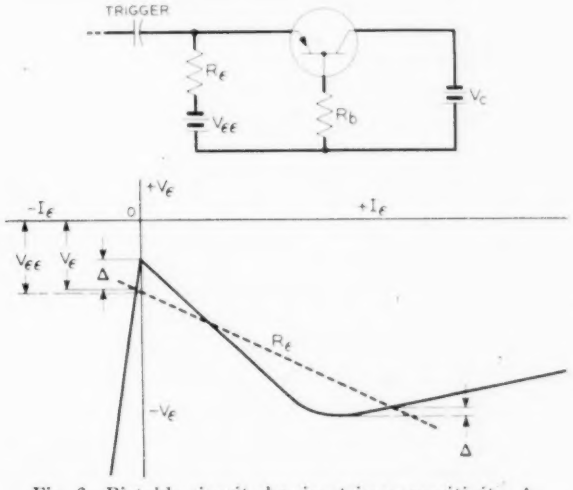


Fig. 6—Bistable circuit showing trigger sensitivity, Δ.

rapidly to a high value of current, falls relatively slowly to an intermediate value, then shifts rapidly to a small negative value and finally returns slowly to the original value. The emitter current and voltage are sketched in Fig. 8. It is a so-called "single-shot" circuit. Alternately the rest or dc stable point can be chosen to be in Region III, at high current, by choice of positive instead of negative bias V_{cc}. Practical considerations as ease in triggering and average power consumption usually indicate a preference for the Region I dc stable point.

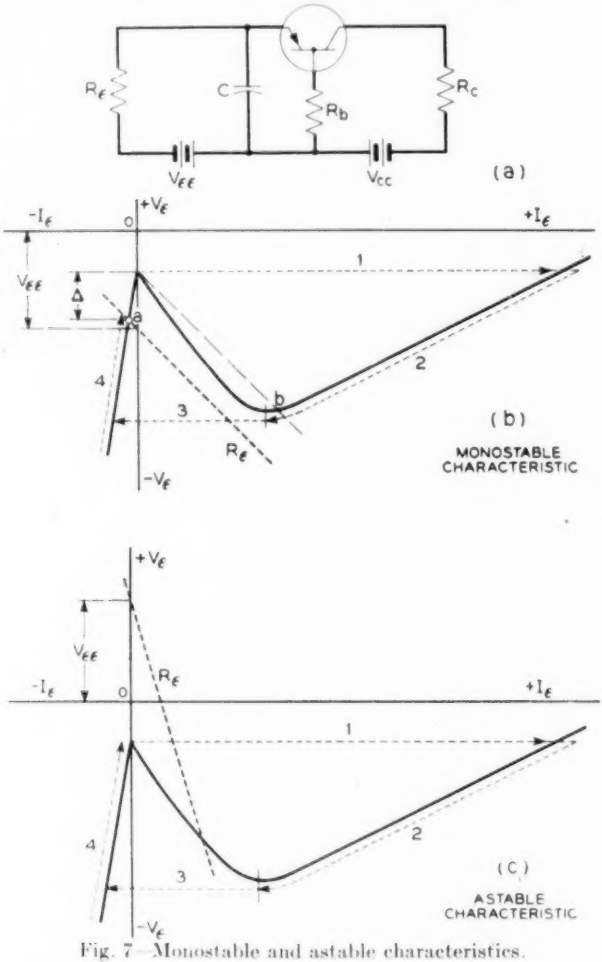


Fig. 7—Monostable and astable characteristics.

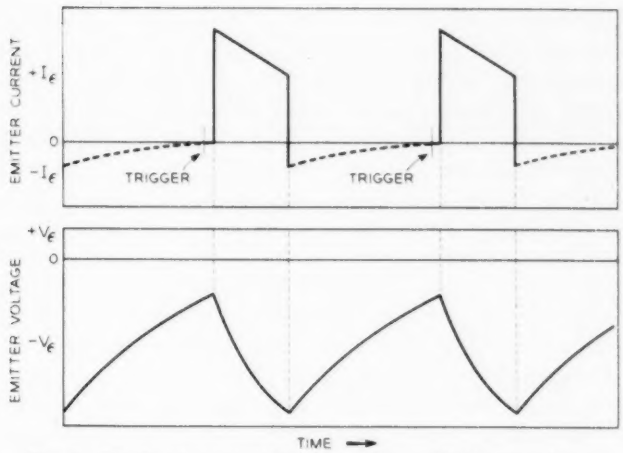


Fig. 8—Idealized monostable relaxation oscillator waveforms.

When the load line and bias are chosen to result in intersection in the negative resistance portion, astable operation or continuous oscillation results. This mode is illustrated in Fig. 7(c). Proper bias and $R_e > |-R_{in}|$, Region II, are required. The operating point formed by the intersection of the load line on the negative resistance portion of the characteristic would normally be stable. However, the capacitor provides an ac short-circuit in parallel with R_e causing the path (1), (2), (3), (4) to be followed continuously. Another form of physical explanation of this relaxation oscillation, usually applied to gas tubes, is that the capacitor C is charged slowly throuch R_e to a critical or breakdown value whereupon the tube or device rapidly discharges the capacitor. When the capacitor charge is dissipated, the device discharge can no longer be maintained due to the IR drop in R_e and the tube or device open-circuits and the capacitor is recharged.

The above suggests a strong simularity to gas tube behavior and this is indeed so. In fact, the modes described above are common to all open-circuit stable negative resistance devices; only the parameters and device phenomena are different.

The primitive circuits of Fig. 7 have properties basic to several switching functions. These may be deduced from the waveforms of Fig. 8 which are essentially identical to both the monostable and astable cases. The emitter current has a rectangular waveform which suggests the generation of rectangular pulses; and, for the astable case, regenerative amplification for both amplitude and wave shape, pulse rate or

frequency division and delay. As shown the current waveform is not particularly good, having neither a flat top nor a flat base line. Practically, the waveform may be derived from the collector by means of a small load resistor to obtain a flat base line. When the emitter current is negative there is sensibly no transfer action, hence, the collector current will be constant during the re-charge portion of the cycle instead of exponential as shown. The slope of the top is inherent and may be removed by clipping. Pulse rise time, the time required for transition from low current to high current, of 0.1 μs is quite easily obtainable; 0.02 μs with average input powers of 20 *mw* have been obtained. Fall time is usually longer than the rise time by factors of 3 or 4. It is to be noted in Fig. 8 that there has been shown a delay between the trigger application and the current transition. Such delay is not peculiar to transistors, but is common to all trigger type devices and circuits. The delay is shown here exaggerated in order to establish its existance and is associated with the static charging of the circuit and the dynamic delay of the device concerned. The trigger-transition delay with transistors is usually less than 0.1 μs.

The voltage waveform of Fig. 8 has a sawtooth form and may thus be employed to generate linear time bases or sweeps. The normal methods for linearization such as a high charging voltage V_{cc} and a high charging resistance R_c or other constant current means are applicable here as in other device circuitry. Free-running and driven sweeps may be obtained with the astable and monostable circuits respectively.

Since the collector characteristic shown in Fig. 3 is also open-circuit stable, the same sort of circuits can be constructed using the output characteristic. Bistable, monostable and astable circuits are shown in Fig. 9.

The resistances seen looking into the base are given in Fig. 10. These circuits are short-circuit stable. That is, high values of R_b result in instability. Bistable, monostable and astable circuits can be constructed also, but use is made of an inductor instead of a capacitor. The reactance of the inductor affords a quasi-open-circuit in the same manner as the capacitor afforded a quasi-short-circuit in the previous cases. Circuit examples are shown in Fig. 11.

SUMMARY

These simple circuits by no means exhaust the switching circuit possibilities of the transistor; rather, they are the simplest. The simple circuit is often satisfactory and may sometimes be employed with little more understanding than that given. More often, however, problems

relating to the sensitivity, constancy of sensitivity, operating currents and voltages, interchangeability and the like require a much more quantitative understanding in order to create circuit designs having specific properties.* An equal need also exists in transistor design for analytic circuit relationships. Such information is useful first, in the

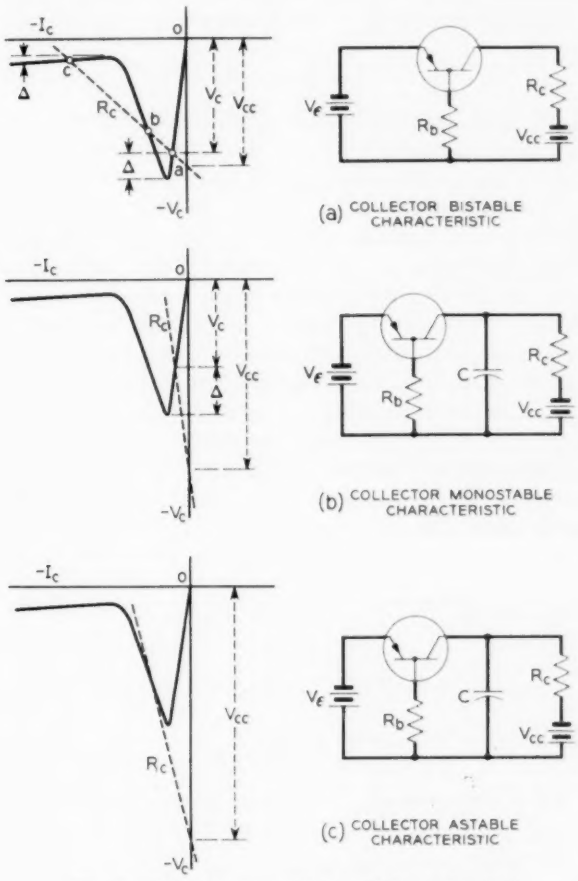


Fig. 9—Collector connection switching circuits.

* See, for example, J. R. Harris, "A Transistor Shift Register and Serial Adder", *Proc. IRE*, Nov., 1952; R. L. Trent, "A Transistor Reversible Binary Counter", *Proc. Inst. Radio Engr.*, Nov., 1952; H. G. Follingstad, J. N. Shive, R. E. Yaeger, "An Optical Position Encoder and Data Transmitter", *Proc. Inst. Radio Engr.*, Nov., 1952.

creation of optimized designs and, second, in the maintainance of proper parameter controls in manufacture. Finally, the more detailed the understanding, the more likely will be the creation of new circuits and new devices.

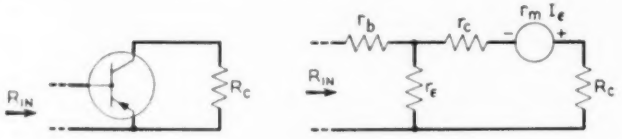
A complete analytical treatment will not be attempted here; consideration will be limited to the equilibrium case and in particular to the simple circuits described.

II—Analysis

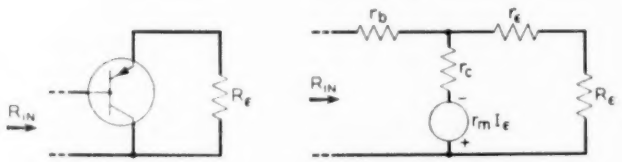
In order to deal analytically with circuits and devices it is necessary to have analytic expressions for the device characteristics. For small signal analysis this is relatively easy. In large signal applications, as in switching, the situation is not so simple. The problems arise because of the high degree of nonlinearity wherein the simplifying assumptions employed in small signal analysis are by no means valid. Further, it is desirable to retain dc terms in many cases.

The method to be employed here is the so called broken-line method which involves approximating the negative resistance characteristic by three intersecting straight lines. The assumption is made that there are three distinct regions of operation in each of which the device is separately linear, but involving different parameter values for each region.

The approximation is shown in Fig. 12. The assumption that the negative resistance characteristic can be simulated by three straight lines is reasonably valid for gross considerations; for fine detail near the



$$R_{IN} = r_b + r_e + \frac{r_e (r_m - r_e)}{R_c + r_e + r_c - r_m}$$



$$R_{IN} = r_b + r_c - r_m + \frac{r_c (r_m - r_c)}{R_e + r_e + r_c - r_m}$$

Fig. 10—Base driving point resistances.

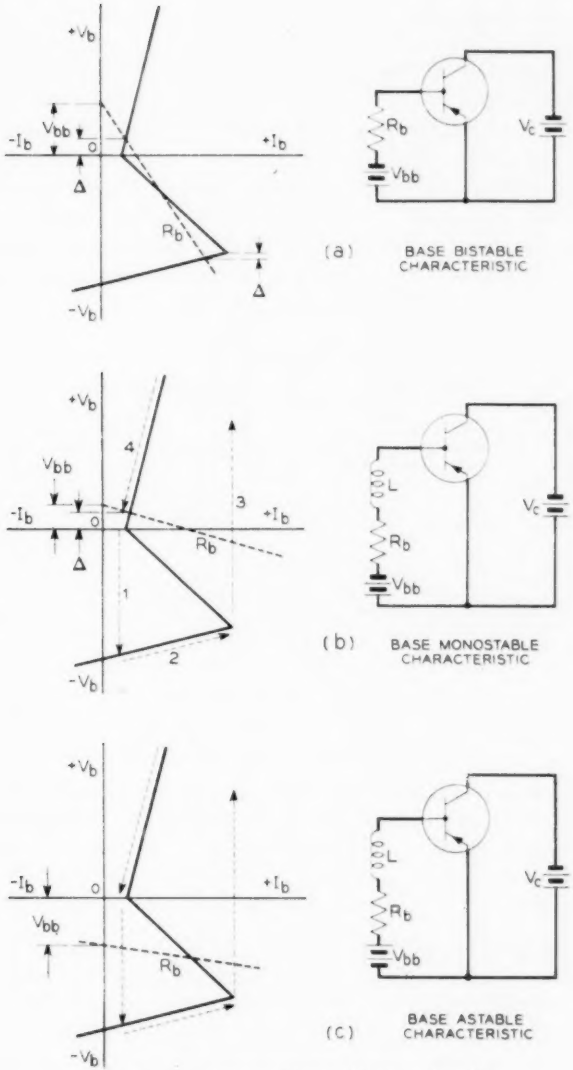


Fig. 11—Base connection switching circuits.

turning points the approximation is by no means accurate although affording zero order information.

Preparatory to the analysis of the negative resistance characteristics, it is necessary to obtain analytic expressions for the transistor currents and voltages. This in turn involves the following steps:

1. Identification of the three regions in terms of the device characteristics,

2. Idealization of the device characteristics to obtains simple, linear relations, and

3. Evaluation of the device parameters in each of the three regions.

Fig. 13 is a family of open circuit characteristics for a typical switching type transistor. Specifically, in small signal terms,

TABLE I

Parameter	*Equivalent Tee*
$R_{11} = \left.\frac{\partial V_e}{\partial I_e}\right]_{I_c}$	$R_{11} = r_e + r_b$
$R_{12} = \left.\frac{\partial V_e}{\partial I_c}\right]_{I_e}$	$R_{12} = r_b$
$R_{21} = \left.\frac{\partial V_c}{\partial I_e}\right]_{I_c}$	$R_{21} = r_m + r_b$
$R_{22} = \left.\frac{\partial V_c}{\partial I_c}\right]_{I_e}$	$R_{22} = r_c + r_b$

Also $$\alpha = -\left.\frac{\partial I_c}{\partial I_e}\right]_{V_c} = \frac{R_{21}}{R_{22}} = \frac{r_m + r_b}{r_c + r_b}$$

The above set, normally employed for small signal analysis, will be assumed to be constant within a given region, but changing in value from region to region.

IDENTIFICATION OF THE THREE REGIONS

It may be recalled with the aid of Fig. 12 that the negative resistance characteristic consists of a negative resistance region bounded on each side by a region of positive resistance. Thus the device is first passive in nature with little or no gain, then very abruptly exhibits considerable gain with the resultant negative resistance, and finally becomes very abruptly passive again with little or no gain.

It would seem quite clear that abrupt changes in the transmission properties of a device should be associated with equally abrupt changes

in the forward transfer characteristic. In the case of the transistor, the behavior of the forward transfer properties is given by the forward transfer impedance, R_{21}.

Examining the R_{21} family in Fig. 13, it is seen that in the normal, positive emitter current region the slope, R_{21}, is high indicating the possibility of high forward gain. When I_e is negative, however, the slope is zero or nearly so, changing very abruptly at $I_e = 0$. Further, it is to be noted that as I_e is made negative, the collector voltage is unaffected, remaining constant for further change in I_e. Thus it may be said that the collector voltage is saturated.*

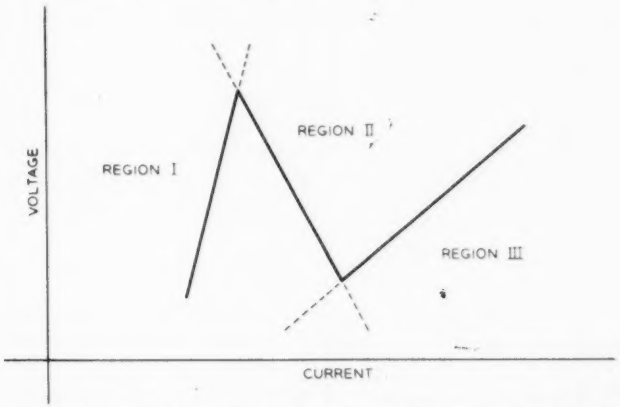


Fig. 12—Broken-line idealization of negative resistance characteristic—division into regions.

If, on the other hand, the emitter current is increased, at constant collector current, it is found that at a critical emitter current the slope again becomes zero or nearly so. There are also two further observations. First, the collector voltage is reduced to a very small value and second, that the critical emitter current is related to the collector current. From the small-signal relation,

$$V_c = R_{21}I_e + R_{22}I_c \tag{2}$$

or

$$V_c = R_{22}\alpha I_e + R_{22}I_c\,, \tag{3}$$

* It is tacitly assumed that in the relation $y = f(x)$ that there are extremes at which y becomes essentially constant and independent of further change in the independent variable x. The point farthest removed from the origin at which the dependent variable becomes constant is termed saturation. The point closest to the origin at which the dependent variable becomes constant is termed cutoff.

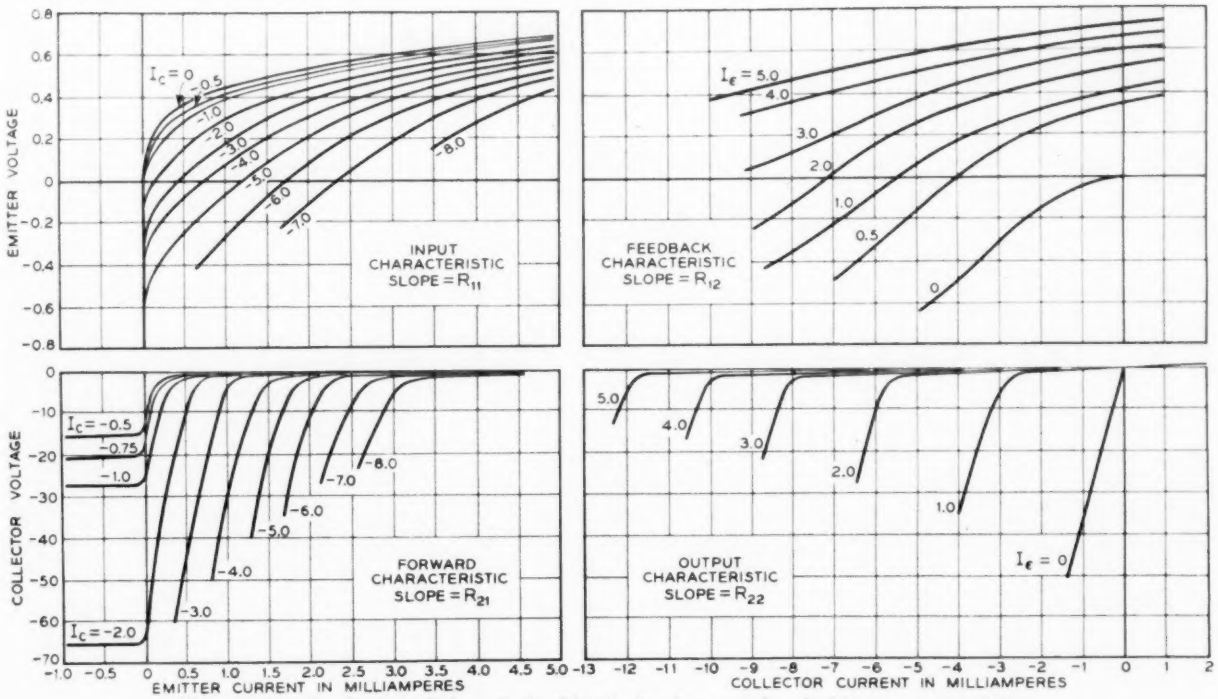


Fig. 13—Static characteristics of the M1689 developmental switching type transistor.

the critical emitter current for collector voltage cutoff may be obtained by setting $V_c = 0$, as,

$$I_e = -\frac{I_c}{\alpha} \tag{4}$$

This relationship is dual to the grid voltage-plate voltage relation in tubes for plate current cutoff as, $V_g = -(V_p/\mu)$. The criteria for defining the three regions are thus established as:

$$\text{Region I (Collector Voltage Saturation): } I_e < 0 \tag{5}$$

$$\text{Region II (Active): } 0 < I_e < -\frac{I_c}{\alpha} \tag{6}$$

$$\text{Region III (Collector Voltage Cutoff): } I_e > -\frac{I_c}{\alpha} \tag{7}$$

The identification of device parameters will be made for the several regions by a single prime for Region I as r_e', none for Region II as r_e, and three primes for Region III as r_e'''.

LINEARIZATION OF THE CHARACTERISTICS AND APPROXIMATIONS

The next step is to linearize the characteristics and to make suitable approximations in order to obtain simple linear equations of the terminal currents and voltages. The relations which require linearization are the device parameters R_{11}, R_{12}, R_{21} and R_{22} which are in general functions of the currents as $R_{ij} = f(I_1, I_2)$.

LINEARIZATION OF R_{11} AND R_{12}

In terms of the equivalent tee circuit, which has been and will be employed, R_{11} is given as $R_{11} = r_e + r_b$. Also, $R_{12} = r_b$. It is convenient to separate r_e and r_b and discuss each separately since r_b is fairly constant and r_e will have widely different regional values.

In the R_{12} family of Fig. 13, it may be seen that R_{12} or r_b is fairly constant in all three regions and will be so taken here. Further, in the simple circuits under consideration, external base resistance R_b has been inserted so that minor variations in r_b in the total of $r_b + R_b$ are inconsequential since usually $R_b \gg r_b$. The approximation that r_b is constant is subject to review where finer detail is necessary, particularly at low emitter currents where the rate of change of r_b is at a maximum.

The emitter resistance r_e is approximately the resistance of a diode in the forward direction. As such, r_e is high when the emitter current is

negative and low when the emitter current is positive. Experimentally, it is found convenient to give three values to r_e and hence to R_{11}, one for each region as shown in Fig. 14. This recognizes the non-linearity with I_e in the forward direction and assumes that a single value in the reverse direction is sufficient. As the circuitry becomes more sophisticated a more precise approximation will undoubtedly be required, particularly near $I_e = 0$.

It may be noted that in the functional relation $R_{11} = f(I_e, I_c)$ that R_{11} is taken to be a function of I_e only. The contribution of I_c is to shift the characteristic in voltage by $r_b\Delta I_c$ increments. Thus the relationship of $V_e = f(I_e, I_c)$ can be written very simply as

$$V_e = R_{11}I_e + R_{12}I_c \tag{8}$$

Since the problem has been linearized to first order terms only, the currents and voltages are total instantaneous or dc values as indicated by the capital letters.

IDEALIZATION OF R_{21}

As indicated previously, R_{21} will be small in Regions I and III and large in Region II. Since $R_{21} = r_m + r_b$, R_{21} can be no less than r_b; the defining approximations will be applied to r_m. In Region I when the emitter current is negative, r_m is taken to be zero and reflects the device approximation that the emitter current under this condition is entirely electron current. This is not always a true approximation, particularly near $I_e = 0$, and limiting tests are employed in transistor testing.

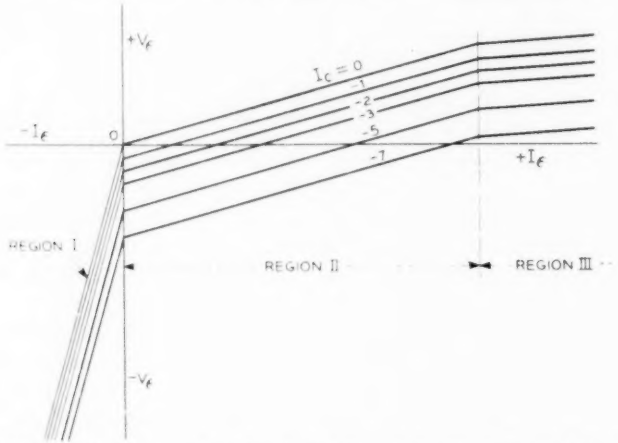


Fig. 14—Idealization and regional division of input characteristic (R_{11}).

In Region II, r_m has, of course, high values and in general $r_m \gg r_b$. Pending further investigation, r_m will be assumed finite, but very small in Region III.

IDEALIZATION OF R_{22}

In the output family of Fig. 13 it may be noted that R_{22} has two values, a high value for $I_e > -\alpha I_e$ and a low value for $I_c < -\alpha I_e$. The high value corresponds to Regions I and II and the low value to Region III. To a first order the two values are separately constant which was not true of earlier transistors in which R_{22} underwent extensive degradation in magnitude as I_c and I_e increased.

The lower limit to which R_{22} can fall in Region III is r_b, since $R_{22} = r_c + r_b$, implying that r_c is zero in Region III. This is approximately, but not accurately true. As αI_e approaches $-I_c$ in magnitude, the voltage across the collector barrier becomes nearly zero so that r_c has a low, but finite value. Under this condition, the hole current is very high and heavy conductivity modulation of the collector barrier resistance occurs. Thus the collector resistance in Region III is indeed quite low and may be neglected for many circuit computations.

In the functional relation $R_{22} = f(I_e, I_c)$ it has been assumed that R_{22} is a function of I_c alone. Further, the approximation involves first order terms only and hence the functional relation $V_c = f(I_e, I_c)$ may be written as:

$$V_c = R_{21}I_e + R_{22}I_c \tag{9}$$

Here again, as in the input case, the currents and voltages are total instantaneous or dc values as indicated by the capital letters.

It is believed desirable, however, to give one more consideration to the output relations. When $I_e = 0$, the collector characteristic is approximately that of a diode in the reverse direction. A diode has low reverse resistance until the voltage across the barrier exceeds a few tenths of a volt and then has quite high resistance, approaching infinite slope in the case of junction diodes.[6, 7] This effect is shown exaggerated in the idealized output family of Fig. 15. The current and voltage at the break in the $I_e = 0$ curve have been termed I_{c0} and V_{c0} respectively. I_{c0} and V_{c0} are quite evident in junction devices; in point contact devices they are not nearly so evident due to the lower value of R_{22} and the higher voltages and currents normally employed. Where currents and

[6] See Reference 2.
[7] *Holes and Electrons*, W. Shockley, Van Nostrand, p. 91, 1950.

voltages are of the order of several milliamperes and a few volts, I_{c0} and V_{c0} may normally be neglected. I_{c0} and V_{c0} do have considerable interest to the device designer, however. The net circuit interpretation of I_{c0} and V_{c0} is to effectively transfer the current-voltage axis from 0, 0 to I_{c0}, V_{c0}. Therefore,

$$V_c - V_{c0} = R_{21}I_e + (I_c - I_{c0})R_{22} \tag{10}$$

or

$$V_c - V_{c0} = (r_m + r_b)I_e + (I_c - I_{c0})(r_c + r_b) \tag{11}$$

Making the approximation that $V_{c0} = I_{c0}R_{22}'''$ and rearranging, equation (10) becomes,

$$V_c - I_{c0}R_{22}''' + I_{c0}R_{22} = R_{21}I_e + I_cR_{22} \tag{12}$$

or

$$V_c + I_{c0}(R_{22} - R_{22}''') = R_{21}I_e + I_cR_{22} \tag{13}$$

which is of the usual form except that a small dc generator of magnitude $I_{c0}(R_{22} - R_{22}''')$ has been added in series with the collector. Since $R_{22} = r_c + r_b$ and $R_{22}''' = r_c''' + r_b$,

$$I_{c0}(R_{22} - R_{22}''') = I_{c0}(r_c - r_c''') \tag{14}$$

The output family equation with equivalent circuit parameters is

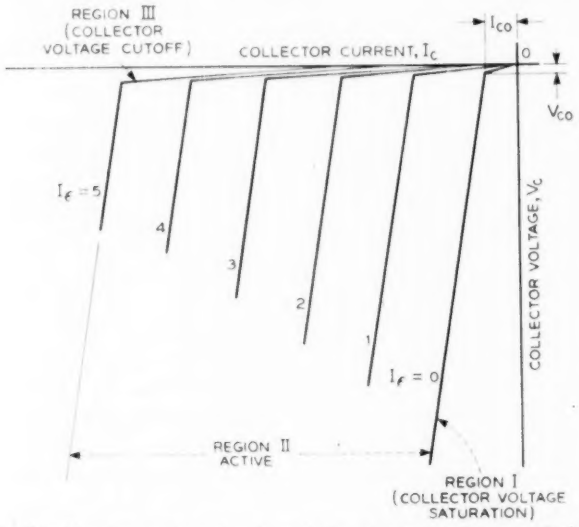


Fig. 15—Idealization and regional division of output characteristic (R_{22}).

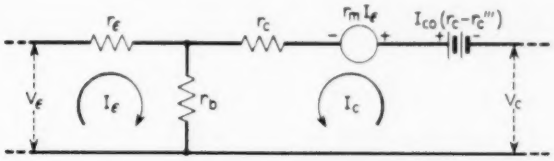
then:

$$V_c + I_{c0}(r_c - r_c''') = (r_m + r_b)I_e + (r_c + r_b)I_c \tag{15}$$

SUMMARY OF IDEALIZATION OF CHARACTERISTICS

The results of the idealization of the device characteristics are summarized in Fig. 16. Here are given analytic expressions for the input and output voltages in terms of the input and output currents; the regions are defined symbolically and by typical values; and an equivalent circuit is given. It may be noted that the equivalent circuit is identical to the small-signal equivalent tee, excepting the small dc generator $I_{c0}(r_c - r_c''')$ which usually may be neglected when dealing with contemporary point contact transistors.

To obtain any of the negative resistance characteristics it is only necessary first to solve the two equations simultaneously for the appropriate voltage in terms of the appropriate current, and then second to insert into the resultant equation the proper parameter values, region by region, to obtain three equations. These equations, when plotted, result in an idealized characteristic similar in form to that of Fig. 12. A detailed example plus synopsis of the properties of the several connections will be given in the following sub-section.



$$V_e = (r_e + r_b)I_e + r_b I_c$$
$$V_c + (r_c - r_c''')I_{c0} = (r_m + r_b)I_e + (r_c + r_b)I_c$$

REGION	PARAMETER							
	r_e		r_b		r_c		r_m	
	SYMBOL	TYPICAL	SYMBOL	TYPICAL	SYMBOL	TYPICAL	SYMBOL	TYPICAL
I	r_e'	100 K	r_b	160	r_c	20 K	r_m'	0
II	r_e	100	r_b	160	r_c	20 K	r_m	50 K
III	r_e'''	25	r_b	50	r_c'''	70	r_m'''	30

$I_{c0} \approx -50\mu A$

Fig. 16—Broken-line transistor equations, regional parameter values and equivalent tee circuit.

THE ALPHA OR CURRENT GAIN FACTOR*

The derivation just given has been in terms of the equivalent circuit parameters, r_e , r_b , r_c , and r_m . Another circuit factor, alpha or the short-circuit current gain, is also quite useful. Alpha has been defined in Table I as the negative ratio of the incremental change in output current to the incremental change in input current *under the condition of short-circuit output terminals.*

Thus alpha is restricted in interpretation to a specific device termination and care should be taken in the employment of alpha when other terminations are involved. For example, the circuit current gain under general conditions is given by R'_{21}/R'_{22} . The ratio R'_{21}/R'_{22} has been sometimes termed α_c . Thus,

$$\alpha_c = \frac{R'_{21}}{R'_{22}} = \frac{r_b + R_b + r_m}{r_b + R_b + r_c + R_c} \tag{16}$$

Depending upon the magnitudes of R_b and R_c , the two current gain ratios may be markedly different. In Region II where r_m and r_c are very large the effects of R_b and R_c in equation (16) often may be neglected. The circuit current gain, α_c , may then be taken as the device alpha. In Region I, r_m has been taken as zero; hence the current gain will be somewhat less than unity, given by $(r_b + R_b)/(r_b + R_b + r_c + R_c)$, and is definitely not zero. Equally, in Region III, the circuit current gain is not zero but rather approaches the ratio, $(r_b + R_b)/(r_b + R_b + R_c)$. If $R_b \gg R_c$, the ratio is nearly unity.

ANALYSIS OF NEGATIVE RESISTANCE CHARACTERISTIC

The objectives of the circuit analysis, as stated previously, are:

1. To determine operating conditions such as proper values of loads, biases, trigger sensitivities and operating currents and voltages,

2. To determine the relationships of the device parameters to the circuit behavior in order that these parameters may be optimized, properly characterized and controlled for required circuit performance.

For example, the trigger sensitivity may be given by the voltage difference between the load line intersection with the Region I portion of the characteristic and the upper peak or turning point of the characteristic as shown in Figs. 6, 7 and 9. The sensitivity Δ is thus determined by the nearness of the bias point to the peak of the characteristic. Since the bias is normally fixed, variations in the sensitivity will arise

* This section is inserted parenthetically as clarifying material due to the use of the α-factor in subsequent discussion.

from variations in the peak point. Thus it is necessary to know the relationships which determine the currents and voltages of the peak and valley points in order first to achieve a design and second, to establish controls on the proper device parameters.

In this example the emitter negative resistance characteristic will be solved and analyzed. The solutions for the other characteristics follow in the same manner and will be summarized.

EVALUATION OF EMITTER CHARACTERISTIC AS AN EXAMPLE

To obtain the emitter characteristic, it is necessary to solve the two equations of Fig. 16 for V_e in terms of I_e. The equations as given are for the short-circuit case. Since the general case will involve external parameters as R_e and R_c, the equations will be modified to include these parameters.

The effects of external parameters may be applied very easily since,

$$V_e = V_{ee} - I_e R_e \tag{17}$$

and

$$V_c = V_{cc} - I_c R_c \tag{18}$$

where V_{ee} and V_{cc} are supply voltages; V_e and V_c are measured from the appropriate terminal to the far end of the external base resistor. External R_b adds directly to r_b. Thus the two equations become:

$$V_{ee} = (r_e + R_e + r_b + R_b)I_e + (r_b + R_b)I_c \tag{19}$$

$$V_{cc} + (r_c - r_e''')I_{c0} = (r_m + r_b + R_b)I_e + (r_b + R_b + r_c + R_c)I_c \tag{20}$$

In manipulation of equations (19) and (20) it is often more easy to do so in the functional form,

$$V_1 = R'_{11}I_1 + R'_{12}I_2 \tag{21}$$

$$V_2 = R'_{21}I_1 + R'_{22}I_2 \tag{22}$$

with substitution at the evaluation stage. The R''s here include both device and circuit parameters.*

Solving equations (19) and (20) simultaneously, the following rela-

* Here the primes indicate generalized open circuit driving point resistance rather than reference to Region I. The duplication of symbols is regretted.

tionship between V_e and I_e is obtained:

$$V_e = \left[r_e + R_e + r_b + R_b - \frac{(r_b + R_b)(r_b + R_b + r_m)}{r_b + R_b + r_c + R_c} \right] I_e + \frac{(V_{cc} + I_{c0}(r_c - r_c'''))}{r_b + R_b + r_c + R_c} (r_b + R_b) \tag{23}$$

Equation (23) is general for the given circuit; it suffers, however, in difficulty in interpretation due to the numerous terms. In the regional evaluation to follow, approximations will be made which bring out the significant factors although decreasing the accuracy somewhat. The $(r_c - r_c''')I_{c0}$ terms will be neglected. It is assumed also that large external base resistance R_b is employed.

EVALUATION IN REGION I

In Region I, from Fig. 16, r_m is zero and r_e' is large so that $r_e' \gg (r_b + R_b)$. Also, by assumption, $r_b \ll R_b$. Applying these approximations, equation (23) becomes,

$$V_{eI} \approx r_e' I_e + \frac{V_{cc} R_b}{R_b + r_c + R_c} \tag{24}$$

Equation (24) is the equation of a straight line, having slope r_e' and an intercept on the voltage axis at $(V_{cc}R_b)/(R_b + r_c + R_c)$. The small-signal input impedance is just the slope value or r_e'.

The short-circuit case where R_c is zero is the most adverse device condition in the sense that the dc term will then be most dependent upon device parameters. When $R_c = 0$, equation (24) becomes

$$V_{eI} \approx r_e' I_e + \frac{V_c R_b}{R_b + r_c} \tag{25}$$

EVALUATION IN REGION II

In Region II all parameters are finite and the only approximations which may be made are $r_b \ll R_b$ and $r_e \ll R_b$. Thus,

$$V_{eII} \approx \left[R_b - \frac{R_b(R_b + r_m)}{R_b + r_m} \right] I_e + \frac{V_{cc} R_b}{R_b + r_c + R_c} \tag{26}$$

If R_b is not too large, it may be approximated that $(R_b + r_m)/(R_b + r_c) = \alpha$. Taking $R_c = 0$, thus,

$$V_{eII} \approx R_b(1 - \alpha) + \frac{V_c R_b}{R_b + r_c} \tag{27}$$

Equation (27) is also the equation of a straight line having the voltage axis intercept of $(V_c R_b)/(R_b + r_c)$ the same value as in Region I. The slope, $R_b(1 - \alpha)$, is negative provided $\alpha > 1$.

EVALUATION IN REGION III

In Region III it may be assumed that $r_b \ll R_b$, $r_c''' \ll R_b$ and $r_m''' \ll R_b$. Other suitable approximations will depend largely upon the magnitude of R_c. From equation (23)

$$V_{eIII} \approx \left[r_e''' + R_b - \frac{R_b(R_b + r_m''')}{R_b + r_c''' + R_c} \right] I_e + \frac{V_{cc} R_b}{R_b + R_c} \tag{28}$$

If R_c is large, that is, large compared to r_e''', but small compared to Region II r_c, then (28) becomes,

$$V_{eIII} \approx \frac{R_b R_c}{R_b + R_c} I_e + \frac{V_{cc} R_b}{R_b + R_c} \tag{29}$$

Under these conditions, the circuit is essentially independent of device parameters. This is useful where a high independence of device parameters is required, but does not focus the attention upon the device parameters as does the $R_c \to 0$ case. This is the condition under which the transistor might be operated when it is desired to obtain the maximum ON current, or conversely the minimum internal switch resistance.

Where $R_c = 0$, equation (28) becomes,

$$V_{eIII} = [r_e''' + r_c''' - r_m''']I_e + V_c \tag{30}$$

Since r_e''' and $(r_c''' - r_m''')$ are quite small the short-circuit currents may be very high. Where the transistor is considered as a switch between emitter and collector circuits, the "switch" voltage drop, as V_{ec}, is given by the first term of equation (30).

EVALUATION OF REGION I–REGION II TRANSITION

Earlier, trigger sensitivities were mentioned as being the small voltage and current differences between the turning points of the negative resistance characteristic and the stable operating points. The determination of the turning points and their stability is of great importance since it is usually desired to obtain maximum stable sensitivity. The voltage and current at the two turning points* have been given the subscript p and v for the low and high current conditions respectively as shown in the synopses of Fig. 17, 18 and 19. V_{ep} and I_{ep} in the short-

* Sometimes termed "peak" and "valley".

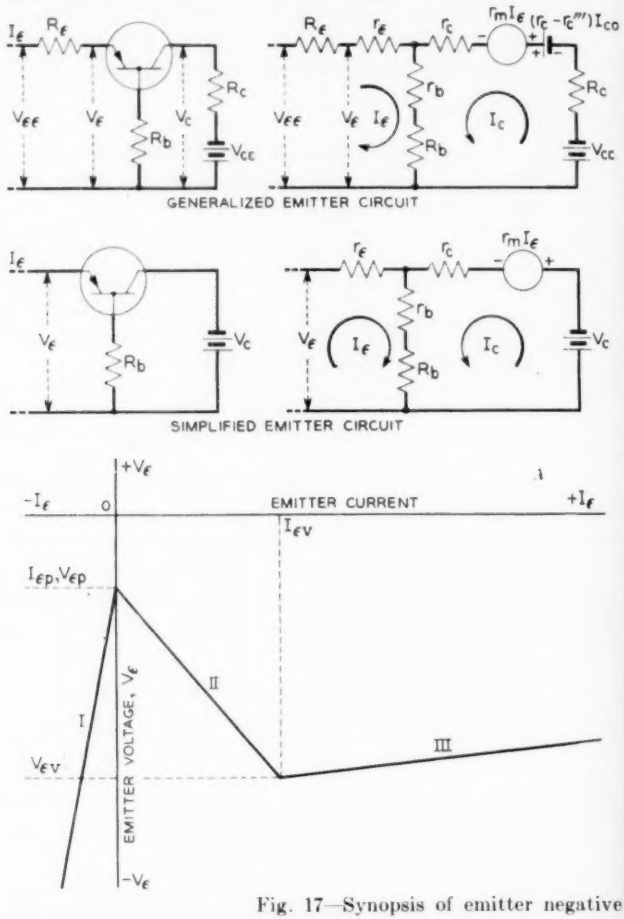


Fig. 17—Synopsis of emitter negative

circuit or $R_\epsilon = R_c = 0$ case for example may be obtained by a simultaneous solution of equation (24) for Region I and equation (27) for Region II. Thus

$$V_{\epsilon p} \approx \frac{V_c R_b}{R_b + r_c} \tag{31}$$

and

$$I_{\epsilon p} = 0. \tag{32}$$

That the low current turning point falls on the emitter current axis, i.e., $I_{\epsilon p} = 0$, is a consequence of the original assumption that $r_m = 0$

Synopsis

General

$$V_{ss} = I_s\left[(r_s + r_b + R_b + R_s) - \frac{(r_b + R_b)(r_b + R_b + r_m)}{r_b + R_b + r_c + R_e}\right] \frac{(V_{cc} + I_{co}(r_c - r_c'''))}{r_b + R_b + r_s + R_c}(r_b + R_b)$$

Approximate Short Circuit Case

where

$$r_b \ll R_b; \quad r_s, r_c''' \ll R_b; \quad R_b \ll r_s, r_m; \quad R_s = R_c = 0; \quad I_{co}(r_c - r_c''') \ll V_c$$

Region I

$$V_s = I_s r_s' + \frac{V_c R_b}{R_b + r_c}$$

Region II

$$V_s = I_s R_b(r - \alpha) + \frac{V_c R_b}{R_b + r_c}$$

Region III

$$V_s = I_s(r_s''' + r_c''' - r_m''') + V_c$$

$$I_{sp} = 0; \quad V_{sp} = \frac{V_c R_b}{R_b + r_c}$$

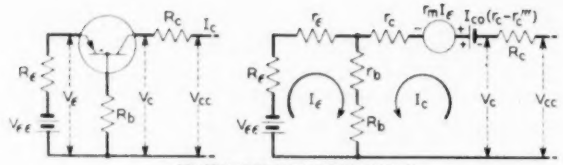
$$I_{ss} = \frac{V_c}{R_b(1 - \alpha)}; \quad V_{ss} = V_c\left[1 + \frac{r_s''' + r_c''' - r_m'''}{R_b(1 - \alpha)}\right]$$

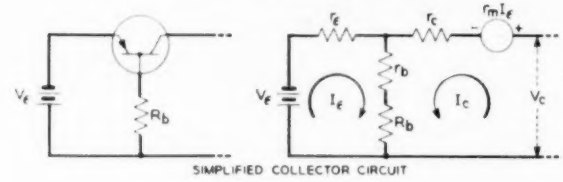
$$\frac{V_{ss}}{V_{sp}} = \frac{R_b + r_c}{R_b}$$

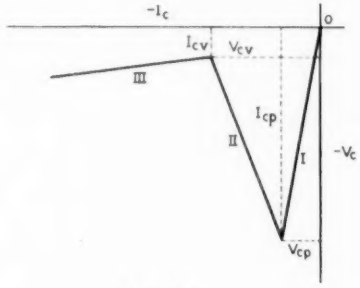
resistance characteristic and properties.

for $I_s < 0$ and $r_m > 0$ for $I_s > 0$. This is not a precise assumption and the turning point will usually lie slightly in the positive emitter current region. For very small triggers or more accurate calculations, consideration must be given to closer approximations of $r_m = f_1(I_s)$ and $R_{11} = f_2(I_s)$.

The consequences of equation (31) can be quite serious. V_c and R_b are of course fixed, but r_c is variable from unit to unit, with temperature and perhaps with life. The variability of V_{sp} can result in failure to trigger, self-triggering or lock-up at high current.







Synopsis

General

$$V_{cc} + I_{co}(r_c - r_c''') = I_c\left[r_c + R_e + r_b + R_b - \frac{(r_b + R_b)(r_b + R_b + r_m)}{r_e + R_e + r_b + R_b}\right] + \frac{V_{ee}(r_b + R_b + r_m)}{r_e + R_e + r_b + R_b}$$

Approximate Short Circuit Case

where

$$R_e = R_e = 0\,; \qquad r_b \ll R_b\,; \qquad I_{co}(r_c - r_c''') \ll V_{cc}\,,\ R_b \ll r_m\,,\ r_c$$

Region I

$$V_e = I_c(r_c + R_b) + V_e\frac{R_b}{r_e'}$$

Region II

$$V_e = I_c[r_c(1 - \alpha)] + \frac{V_e(r_m + R_b)}{R_b}$$

Region III

$$V_e = I_e(r_e''' + r_c''' - r_m''') + V_e$$

$$V_{cp} = V_e\left(\frac{r_c + R_b}{R_b}\right) \qquad I_{cv} = \frac{V_e}{R_b}\left(\frac{\alpha}{\alpha - 1}\right)$$

$$I_{cp} = \frac{V_e}{R_b} \qquad \frac{V_{cp}}{V_{cv}} = \frac{r_c + R_b}{R_b}$$

$$V_{cv} = V_e\left[1 + \frac{\alpha(r_e''' + r_c''' - r_m''')}{R_b(\alpha - 1)}\right]$$

Fig. 18—Synopsis of collector negative resistance characteristic and properties.

EVALUATION OF REGION II–REGION III TRANSITION

The high-current turning point for the short-circuit case is determined from a simultaneous solution of the pertinent equations for Regions II and III, equations (27) and (30). Thus,

$$I_{ev} \approx \frac{V_c r_e}{R_b(1-\alpha)(r_e + R_b)} \tag{33}$$

$$V_{ev} \approx V_c\left[1 + \frac{r_c(r_e''' + r_e''' - r_m''')}{(r_c + R_b)R_b(1-\alpha)}\right] \tag{34}$$

Where it may be approximated that $r_c \gg R_b$, as has already been done in bringing in α, equations (33) and (34) become,

$$I_{ev} \approx \frac{V_c}{R_b(1-\alpha)} \tag{35}$$

$$V_{ev} \approx V_c\left[1 + \frac{r_e''' + r_c''' - r_m'''}{R_b(1-\alpha)}\right] \tag{36}$$

In this order of approximation, V_{ev} is nearly equal to V_c. Any variation in the lower trigger point is primarily with I_{ev}, due chiefly to any change in α. It is interesting to note that the trigger point will move along the Region III curve given by (30).

The ratio of V_{ev} to V_{ep} is often of interest to estimate voltage swings or perhaps as a design objective in some switching circuits. Thus from (36) and (31),

$$\frac{V_{ev}}{V_{ep}} = \frac{[R_b(1-\alpha) + r_e''' + r_c''' - r_m'''](R_b + r_c)}{R_b^2(1-\alpha)} \tag{37}$$

If $r''' + r_c''' - r_m''' \ll R_b(1-\alpha)$ then (37) becomes:

$$\frac{V_{ev}}{V_{ep}} = \frac{R_b + r_c}{R_b} \tag{38}$$

If R_b is very large, the ratio approaches unity with the implication of the existence of only two regions. This is equivalent to saying that the negative resistance becomes infinite over an infinitely short range. The proper choice of R_b in terms of (38) may well be a design problem where it is desired to have a high ratio of V_{ep} to V_{ev}, as in lockout circuits.

SYNOPSIS OF NEGATIVE RESISTANCE CHARACTERISTICS

Synopsis for the three negative resistance characteristics are given in Figs. 17, 18 and 19. The solution and analysis procedures are the same as outlined for the emitter characteristic. It should be noted that the base characteristic is short-circuit stable in distinction to the emitter

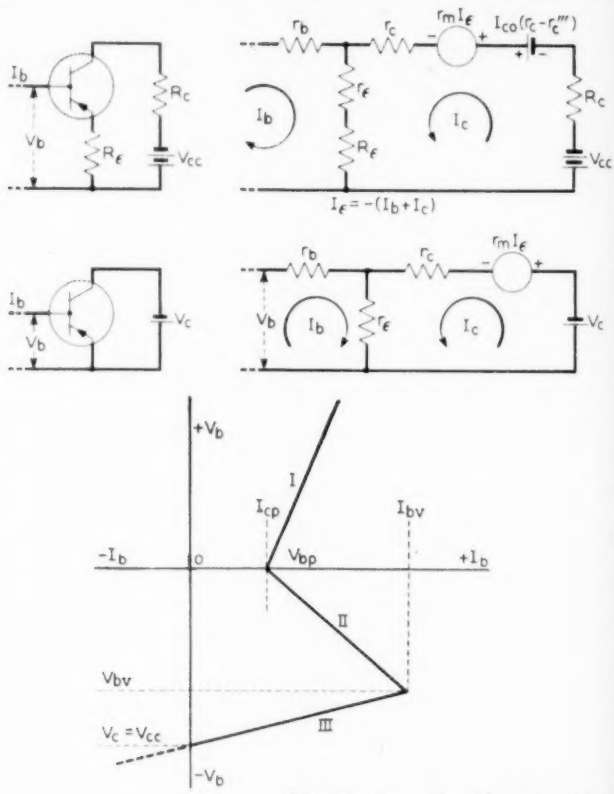


Fig. 19—Synopsis of base negative

and collector characteristics which are open-circuit stable. It would have been more appropriate to solve the base circuit in terms of conductances rather than resistances. The magnitudes of negative resistance obtained in this connection are quite low which may be misleading; the negative conductance is quite high, however, which is desired in short-circuit stable negative resistance circuits.

Care should be taken in the literal employment of the approximate regional relationships in Figs. 17, 18 and 19. They are very definitely approximate and are intended to illustrate behavior and the limiting condition only to bring out the relative importance of device parameters. It is suggested that calculations be started with the general case and approximations be made as are valid. For example, the con-

Synopsis

General

$$V_b = I_b(r_b + R_b + r_e + R_e) + I_c(r_e + R_e)$$

$$V_{cc} + I_{co}(r_c - r_c''') = I_b(r_e + R_e - r_m) + I_c(r_c + R_e + R_g - r_m)$$

$$V_b = I_b\left[r_b + R_b + R_e + r_e - \frac{(r_e + R_e)(r_e + R_e - r_m)}{r_e + R_e + r_c + R_c - r_m}\right] + \frac{[V_{cc} + I_{co}(r_c - r_c''')](r_e + R_e)}{r_e + R_e + r_c + R_c - r_m}$$

Approximate Short Circuit Case

$$R_e = R_c = 0\ ; \quad I_{co}(r_c - r_c''') \ll V_c\ ; \quad r_e \ll r_c(1 - \alpha)$$

Region I

$$V_b = I_b\left(\frac{r_e' r_c}{r_c + r_e'}\right) + \frac{V_c r_e'}{r_e' + r_c}$$

Region II

$$V_b = I_b\left(\frac{r_b + r_e}{1 - \alpha}\right) + \frac{V_c r_e}{r_c(1 - \alpha)}$$

Region III

$$V_b = I_b r_b''' + V_c$$

$$I_{bp} = \frac{V_c}{r_c}\ ; \quad V_{bp} = 0$$

$$I_{bv} = V_c\left[\frac{1 - \alpha}{r_e}\right]\ ; \quad V_{bv} = V_c\left(1 - \frac{(\alpha - 1)r_b'''}{r_e}\right)$$

resistance characteristic and properties.

clusion is reached in the collector characteristic that the negative resistance (Region II) is independent of the base resistance or feedback. This is true for only the limited range where $r_e \ll R_b \ll r_c$.

EXAMPLE OF CALCULATED AND EXPERIMENTAL CHARACTERISTICS

An example to illustrate the analysis is shown in Fig. 20 where both experimental and calculated characteristics for the emitter circuit are given. In this example there is appreciable load resistance; hence r_c''', r_e''' and r_m''' are of no consequence since they will all be very small compared to the R_c of $2.2K$ ohms. Also, $R_b = 6.8K$ ohms is much greater than r_b; hence r_b can be neglected. Since V_c is -45 volts, the I_{co} term may also be neglected.

Computing V_{ep} first,

$$V_{ep} = \frac{V_c R_b}{R_b + r_c + R_c} = \frac{-45(6.8K)}{(6.8K + 19K = 2.2K)} = -10.9 \text{ volts} \quad (39)$$

The calculated value of -10.9 volts compares quite favorably to the measured -11.0 volts.

Region II is given in this case, approximately by,

$$V_e \approx \left[R_b + \frac{R_b(R_b + r_m)}{R_b + r_c + R_c} \right] I_e + \frac{V_c R_b}{R_b + r_c + R_c} \quad (40)$$

or

$$V_e \approx \left(6.8K + \frac{6.8K(6.8K + 50K)}{6.8K + 19K + 2.2K} \right) I_e - 10.9 \quad (41)$$

$$V_e \approx (-8.9K) I_e - 10.9 \quad (42)$$

The first term is of course the slope in Region II and is the magnitude of the negative resistance. The calculated value is -8900 ohms whereas the measured value was approximately -9200 ohms.

The Region III approximation, derived also from the general relationship is,

$$V_e = \frac{(R_b R_c)}{R_b + R_c} I_e + \frac{V_c R_b}{R_b + R_c} \quad (43)$$

$$= \frac{((6.8K)(2.2K))}{(6.8K + 2.2K)} I_e = \frac{45(6.8K)}{6.8K + 2.2K} \quad (44)$$

or

$$V_{eIII} = (1785) I_e - 34 \quad (45)$$

The relation for Region III agrees quite well in slope but not in dc value as may be seen in Fig. 20. Since in this example the Region III behavior is determined essentially by the circuit parameters, it is surmised that the nominal 45-volt battery employed in taking the data was actually 47 volts.

The Region I check is essentially perfect since the approximation given in Fig. 17 is quite good.

Note the error at the intersection of the Regions II and III. The broken-line method predicts a sharp transition whereas the actual case is gradual. The deviation is due to the gradual changes in r_m and r_c as the collector voltage approaches cutoff and is the largest gross error in the approximation.

It is believed that analysis of this sort will reasonably predict circuit behavior and lead to device requirements. There must be a thorough understanding of the approximations involved and the accuracy will be directly related to the degree to which the original idealized characteristics are approximated. Extended, by means of more than three broken

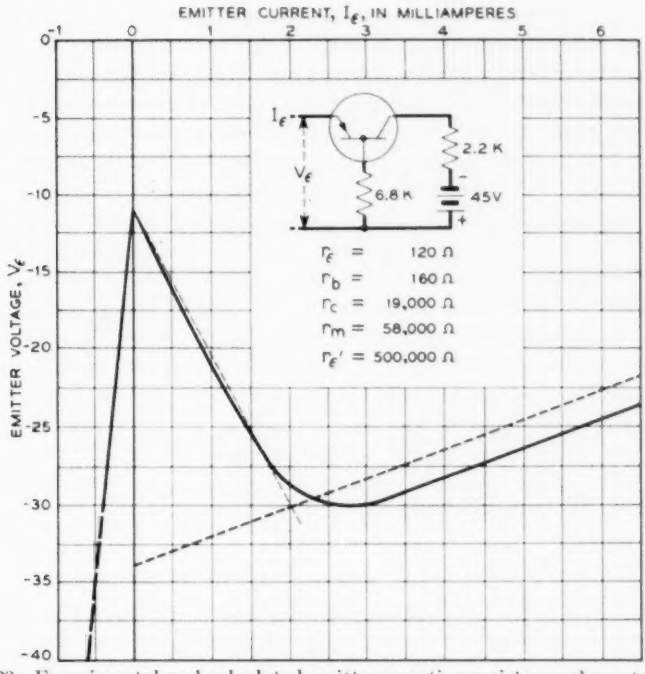


Fig. 20—Experimental and calculated emitter negative resistance characteristic.

lines, the method will yield fine detail to the degree to which device parameters are known and patience will permit. Transient behavior and analysis have not been discussed and are needed for a more complete understanding, particularly where transitional speeds are of concern.*

III—Switching Type Transistor Properties

An examination of the circuit approximations given in Figs. 17, 18, and 19 will reveal that the transistor and circuit designers will want to know nearly all there is to know about the device characteristics. This is not particularly surprising since the device is used over its entire range rather than over a limited portion as in the case of small-signal applications. The same examination of the circuit relations will also show that

* A treatment of the transient behavior between regions is given in B. G. Farley, "Dynamics of Transistor Negative Resistance", *Proc. Inst. Radio Engr.*, Nov., 1952. Analysis and the solution for the periods of the monostable and astable cases, assuming infinite region to region transition speed, are given in G. E. McDuffie, Jr., "Pulse Duration and Repetition Rate of a Transistor Multivibrator", *Proc. Inst. Radio Engr.*, Nov., 1952.

virtually all of the device parameters should be constant from unit to unit and with ambient conditions.

It can be shown that for small-signals a device may be uniquely characterized by five measurements. In terms of the parameters used here these might be R_{11} , R_{12} , R_{22} , R_{21} and the dc bias point or equally, r_e , r_b , r_c , r_m and the bias point. Since the problem was linearized in the approximation, it follows that 15 such measurements, five in each region, are necessary for proper switching device characterization. The indicated extensive testing required may be reduced somewhat by suitable approximations. It is clear that the switching device designer and producer must reconcile themselves to making more tests for accurate characterization than when small-signal devices are concerned.

What will be given here is a description of typical developmental switching transistors in terms of the parameters which have evolved as a result of practical approximations. The method will be to discuss device properties and measurements region by region; then to discuss the properties at the transition points. Temperature, frequency and life behavior will be taken up separately.

REGION I PROPERTIES

In Region I, the emitter current is negative. Hence the emitter resistance r_e' is large and is essentially that of a diode in the reverse direction. At present r_e' is measured by a simple dc test of the current which flows at a nominal -10 volts. Both r_e' and r_b will be discussed further under the Region I–Region II transition properties.

The Region I collector resistance is one of the most important parameters in switching. This is because of its determining nature in the turning point voltages in Figs. 17, 18, and 19. Actually, what is of concern is not the small-signal slope shown as r_c* in Fig. 21, but rather the dc current and voltage relationship shown as r_{c0} . For example in Fig. 17, it may be seen that V_{ep} is given by the voltage drop determined by the product of R_b and the dc collector current.

Fig. 21 is an idealization of the R_{22} characteristic and has been designed to bring out the diode nature of the collector by emphasizing the saturation current and voltage, I_{c0} and V_{c0} . In junction devices the break in the $I_e = 0$ characteristic at I_{c0} is quite evident whereas in present point contact devices the transition is smooth due to the much lower values of r_c . The device significance is the same, however; I_{c0} varies rapidly with temperature whereas r_c varies at a considerably lower rate.

* The actual parameter is of course R_{22} , but since $R_{22} = r_e + r_b$ and $r_b \ll r_c$ R_{22} is taken as r_e .

In junction devices the proper measurements would be of I_{c0} and r_e. Since I_{c0} is difficult to define in point contact devices, r_{c0} has been measured as an approximation. In the idealization, r_e and r_{c0} are related as,

$$r_{c0} \doteq \frac{(I_e - I_{c0})}{I_e} r_e \tag{46}$$

The measurement of r_{c0} is made at a collector voltage which is typical of the applications in the range of perhaps -10 to -45 volts.

A constant dissipation line has been drawn on Fig. 21, which reveals the desirability of having r_{c0} very large in order to operate at higher voltages and to secure high efficiency through lower dissipation in the OFF or rest condition.

REGION II PROPERTIES

The Region II low frequency properties are essentially identical to those of transistors intended for small-signal applications. A possible exception is the somewhat less attention paid to the base resistance, r_b, which is critical to small-signal applications. The characterization consists of a normal small-signal set plus dc bias values.

REGION III PROPERTIES

The Region III properties have been defined largely by a figure of

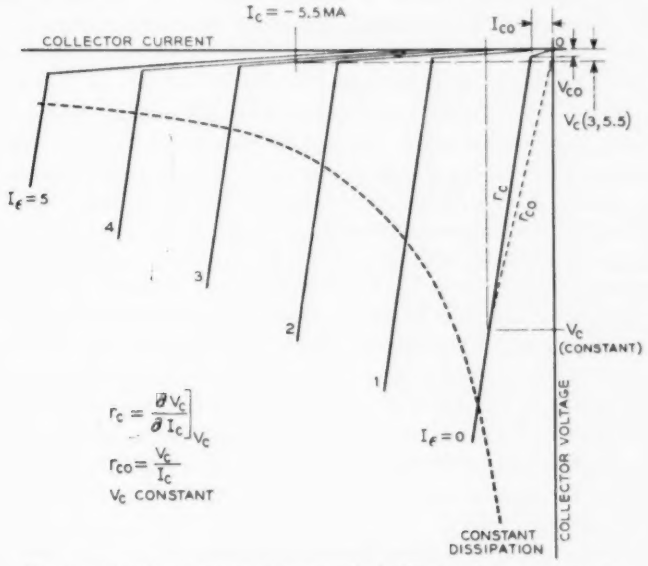


Fig. 21—Idealized output characteristic illustrating parameters.

merit measurement shown as $V_c(3, -5.5)$ in Fig. 21. This measurement is the voltage from collector to base under the condition that $I_e > -(I_c/\alpha)$. In this instance I_e has been chosen to be 3 mA and I_c to be -5.5 mA. The collector current value is chosen on the basis of the smallest tolerable value of alpha expected so as to place the point of measurement near the R_{22} knee, but in Region III or overload.

The $V_c(3, -5.5)$ measurement is a good measurement for defining the general behavior. $V_c(3, -5.5)$ taken with the r_{c0} measurement constitute a very good defining set for checking the transistor as in re-measuring. For design purposes, the $V_c(3, -5.5)$ measurement is not sufficient. It provides an approximate value for r_c''', but does not define r_e'' and r_m'''. A second dc measurement, the collector to emitter voltage drop, V_{ec}, has been employed experimentally also. An improved characterization will undoubtedly involve separate measurements of r_e''', r_m''' and r_c'''.

REGION-TO-REGION TRANSITION PROPERTIES

The transition between Regions I and II is accompanied by abrupt changes in r_e and r_m.

The theory assumes that both of these parameters change at an infinite rate at a fixed emitter current, taken as $I_e = 0$. Unfortunately neither of these assumptions is strictly true. r_e undergoes a gradual change from high to low values which is only approximated by the three assigned values. In particular the behavior near $I_e = 0$ is of concern when dealing with small triggers.

The forward transfer impedance changes at a finite rate also. Further, the emitter current at which the maximum rate of change occurs will vary from unit to unit. Present practice also has been to measure α rather than r_m. The rational for doing so is not too good since r_m is quite likely the better parameter to characterize. Alpha has a strong physical appeal, fits well into the circuit problems and is easy to measure.

Since $\alpha = (r_b + r_m)/(r_b + r_c)$ it is necessary to assume that r_b and r_c are constant near $I_e = 0$, an only fair approximation. Having made the approximation, the typical α behavior shown in Fig. 22 may be taken as a measure of r_m. Three values are measured, the first of which, α_1, in Region II, is redundant to the Region II small-signal measurements. The two limits, α_2 and α_3, serve to place lower and upper limits on the absolute values of α at the Regions I–II transition. These limits in turn place a lower value on the rate of change in α within the $I_e \pm \Delta$ range shown.

It may be noted that α in Region I is finite. There is a lower limit

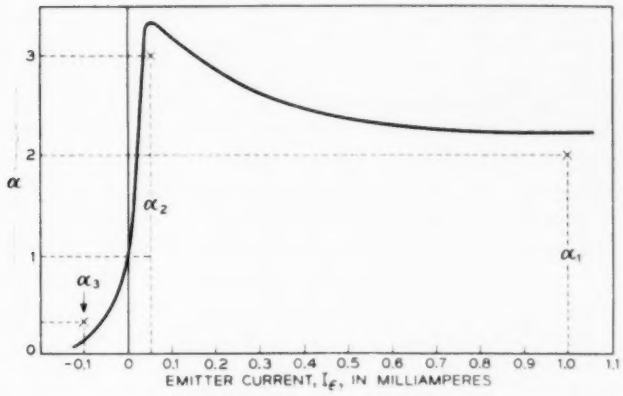


Fig. 22—Alpha characteristic.

even though r_m is zero since $\alpha_1 \rightarrow (r_b/r_b + r_c)$. The values normally encountered at $I_e = 0 - \Delta$ are usually in excess of this lower limit.

The feedback resistance r_b tends to rise as $I_e \rightarrow 0$ which may be important to some trigger circuits. As the circuitry becomes more sophisticated, it is expected that more attention will need to be paid to the behavior of r_e, r_m and r_b at and near $I_e = 0$.

The transition from Region II to Region III is determined from the relation $I_e = -(I_c/\alpha)$. The problem is quite similar to the control of the μ factor in tubes where plate current cut-off is given by $V_g = -(V_p/\mu)$. Present practice has been to depend upon the α_1 values and upon the lower limit placed on alpha in the $V_c(3, -5.5)$ measurement. Further effort is needed here also.

TYPICAL PARAMETER VALUES AND DISTRIBUTIONS

Integrated distribution curves for the parameters of a typical developmental switching transistor are shown in Fig. 23. The unit-to-unit variations are deemed to compare favorably with those of commercial electron tubes. The parameter of most serious variability is r_{c0} which is unfortunate since r_{c0} is so important to trigger sensitivity stability.

TEMPERATURE, FREQUENCY AND SHOCK PROPERTIES

Transistor parameters are reasonably constant with temperatures below room temperature. Above room temperatures some of the parameters are variable. r_e and r_b are fairly constant, changing very little to 70°C. r_c and r_m decrease fairly rapidly, maintaining a ratio such that alpha rises slightly. r_e' and r_{c0} change most rapidly and, while both of

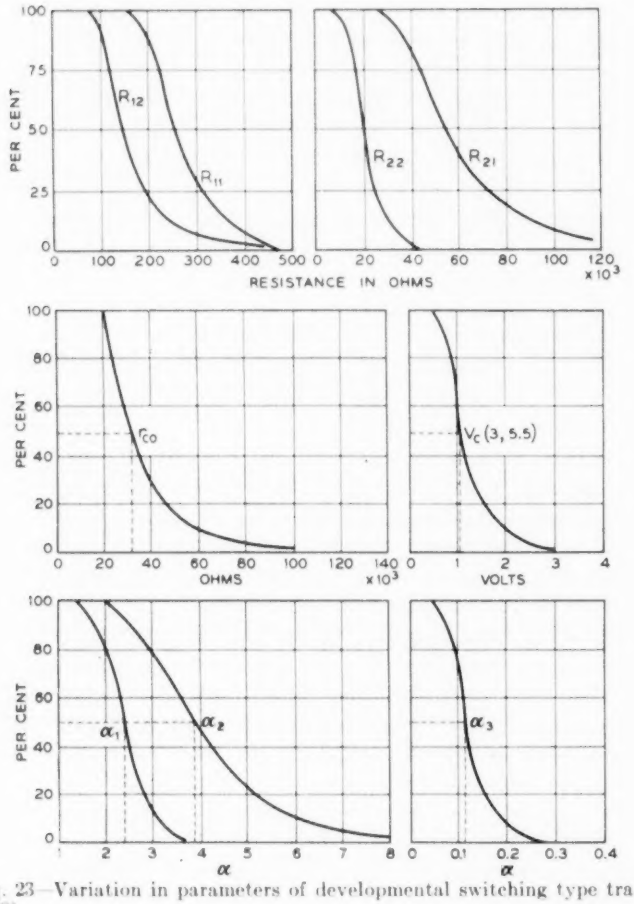


Fig. 23—Variation in parameters of developmental switching type transistor (M1698).

these parameters are of little consequence in small-signal applications, they are quite important in switching, particularly r_{c0}.

Early transistors might exhibit a change in r_{c0} at 60°C of 3 to 1 or more from room temperature values. The transistors of which the data in Figs. 13 and 23 are typical have an r_{c0} temperature coefficient of about −0.75 per cent/°C. That is, the room temperature value of r_{c0} might be reduced by 30 per cent at 70°C. The improved temperature behavior implies a corresponding reduction in variation in trigger sensitivity. Parameter values, large-signal and small-signal, are shown in Fig. 24 as a function of temperature.

Variation in characteristics will arise from self-engendered heat, that is, dissipation. Transistors may be thermally unstable under constant voltage conditions. Since the switching properties are exhibited under short-circuit or constant voltage terminations, thermal properties are of concern. The limitations involved are similar to those of any positive feedback circuit. If the thermal loss through radiation and conduction exceeds the heat input, the system will be stable. The practical significance is to place limitations on dissipation and to employ designs which result in rapid heat loss. Other design criteria such as miniaturization may limit the latter.

If perfect switching characteristics were obtainable, dissipation would be of little consequence in switching. This is akin to saying that neither a short-circuit nor an open-circuit dissipates any energy. Further, the perfect device has zero transition time and therefore involves no loss. The transistor has finite resistance both open and closed and a finite although rapid transition time. There is some advantage however. A constant dissipation curve shown as a dotted line has been included in Fig. 21. Small-signal operation at mid-range currents and voltages results in fairly low limitations on both current and voltage. The intersection with the R_{22} voltage saturation line ($I_e = 0$) is at fairly high voltage. Similarly, the intersection with the collector voltage cut off line is at high current. For constant dissipation, approximately,

Voltage saturation: $$P_d \approx \frac{V_c^2}{r_{c0}}$$

Voltage cutoff: $$P_d \approx I_c^2(r_e''' + r_c''' - r_m''')^*$$

Depending upon the circuit the assumed dissipation limit may or may not be exceeded during the transitions. Should the limit be exceeded,

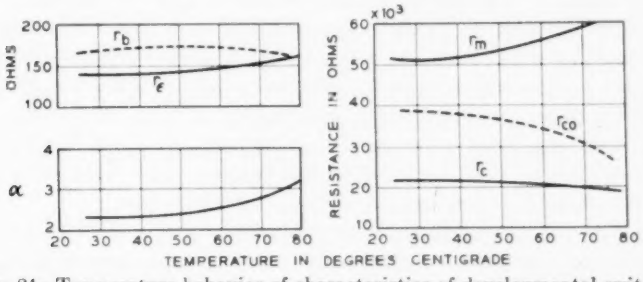


Fig. 24—Temperature behavior of characteristics of developmental switching-type transistor (M1689).

* This includes both emitter and collector dissipation. See equation (30).

and substantially so, there are normally no serious consequences due to the very rapid transitions and consequent low thermal energy generated.

Transistors may not be able to tolerate excess dissipation on this basis if the circuits are slow, that is with transition times in excess of perhaps a few tenths of a microsecond. Such conditions may arise, for example, if loads are inductive. In many such cases, shunting capacitor networks will often permit a rapid transition with consequent transfer of current to the inductive load.

The frequency response of point contact transistors can be sufficiently good to insure switching type operation with rise times of the order of 0.1 to 0.01 μs. Fall times may be somewhat longer due to the hole storage effect. In regenerative circuits, operating speeds are faster than might be imagined from the small-signal frequency cutoff. Reliable operation with rise times of 0.1 μs is obtained with only nominal attention to frequency cutoff. Speeds of the order of 0.02 μs require a 10 mc. lower limit. Present junction transistors are substantially slower.

Accurate life estimates are difficult to make due to the rapid rate of development, the relative age of the transistor and the number of parameters involved. A given device is quite likely to be obsolete and forced to give way to an improved version before sufficient models can be obtained for life tests. A small quantity of transistors having properties similar to those of Fig. 20 and 21 have been operated for over 6,000 hours with an indicated life of 30,000 hours. Other similar transistors with longer life histories have indicated lives of better than 70,000 hours. The pattern appears to be similar to that of electron tubes—an early failure and change rate followed by a very slow exponential rate. It is believed that life is extended by low power operation and is decreased by high temperature operation.

The relatively high noise level of transistors does not appear to be a significant problem at present when considered in terms of automata. Systems employing switching type circuits in pulse communication will of course be concerned. It is suggested that the non-concern for noise in non-transmission type systems is largely a reflection of the ease with which high magnitudes of state changes are obtained. With design trends toward low power and low operating levels, noise will undoubtedly set a lower limit of level operation in such systems also.

The extreme resistance of the transistor to shock and vibration with a consequent absence of microphonism may in some applications result in effective lower noise. Shocks in excess of 20,000G have resulted in no damage. No evidences of current modulation in excess of noise have been detected with vibrational forces of the order of 100G at frequencies

as high as 1000 cycles in tests on the transistor of Fig. 1. Transistors have been included in plastic embedded circuits without change of characteristics.

SUMMARY—TRANSISTOR PROPERTIES

Transistors have been designed with properties expressly intended for switching applications.* The characteristics are acceptable for contemporary switching type circuits and sufficiently reproducible to permit interchangeability of devices in circuits of normal requirements. The characterization has been sufficiently unique to permit the calculation of first order circuit performance. The characterization is not sufficiently complete to permit determination of the complete transient behavior.

In terms of the circuits described, the major parameter limitation is concerned with the variability of the d-c collector resistance among units and with temperature. It is expected that future circuit development will place additional requirements on the transistor, particularly as related to the transitions between regions. It is also to be expected that future circuit designs may establish new or modify present device requirements.

A major consideration for computer or computer-like systems, reliability, particularly with respect to time and temperature, has not been established, but appears to be favorable.

ACKNOWLEDGMENT

It is impossible to properly acknowledge credit to all of those who contributed to the concepts, data and results of this paper. Particular acknowledgment is due to J. A. Morton who provided first the method of attack for the analysis and second, continued stimulation. Acknowledgment is also given to A. J. Rack who first classified and explained the several simple circuits of the first section. J. J. Kleimack provided transistor data and R. L. Trent, circuit data.

* See Reference 3 also.

Abstracts of Bell System Technical Papers* Not Published in This Journal

An Approximate Quantum Theory of the Antiferromagnetic Ground State. P. W. ANDERSON[1]. *Phys. Rev.*, **86,** pp. 694–701, June 1, 1952. (Monograph 1995).

A careful treatment of the zero-point energy of the spin-waves in the Kramers-Heller semiclassical theory of ferromagnetics leads to surprisingly exact results for the properties of the ground state, as shown by Klein and Smith. An analogous treatment of the antiferromagnetic ground state, whose properties were unknown, is here carried out and justified. The results are expected to be valid to order 1/S or better, where S is the spin quantum number of the separate atoms.

The energy of the ground state is computed and found to lie within limits found elsewhere on rigorous grounds. For the linear chain, there is no long-range order in the ground state; for the simple cubic and plane square lattices, a finite long-range order in the ground state is found. The fact that this order can be observed experimentally, somewhat puzzling since one knows the ground state to be a singlet, is explained.

Method of Synthesis of the Statistical and Impact Theories of Pressure Broadening. P. W. ANDERSON[1]. Letter to the Editor. *Phys. Rev.*, **86,** p. 809, June 1, 1952.

Arcing at Electrical Contacts on Closure. Part III. Development of an Arc. L. H. GERMER[1] and J. L. SMITH[1]. *Jl. Applied Phys.*, **23,** pp. 553–562, May, 1952. (Monograph 2002).

A description is given of a system made up of experimental electrodes and an oscilloscope by means of which the potential across the electrodes can be recorded with a time resolution of about 10^{-9} sec. and a potential sensitivity of 1-trace width per volt. The closure of the electrodes to produce a short arc is synchronized with the oscilloscope sweep so that the beginning of the arc is photographed.

As an arc starts the potential across the electrodes decreases more or less gradually from the applied voltage to a steady value characteristic of the metal of the electrodes. The course of this change is extremely variable as is also the time over which the change is spread. The average value of the time appears to

* Certain of these papers are available as Bell System Monographs and may be obtained on request to the Publication Department, Bell Telephone Laboratories, Inc., 463 West Street, New York 14, N. Y. For papers available in this form, the monograph number is given in parentheses following the date of publication, and this number should be given in all requests.

[1] Bell Telephone Laboratories

vary with circuit inductance and with the nature of the electrode surfaces. For inactive silver surfaces and an inductance of 0.10 μh the average value of the time is about 0.007 μ sec., and for active surfaces and the same inductance 0.011 μ sec. For active surfaces and an inductance of 5 μh the average value of the time is 0.02 μ sec.

The electrode separation at which an arc strikes is determined from the oscilloscope traces and from a correction for the height of the mound of metal thrown up by the arc. For active silver electrodes the average separation (at 40 or 45 volts) corresponds to a gross electric field of 0.8×10^6 volts/cm. and for inactive silver electrodes to a field of 2.3×10^6 volts/cm. These are probably better values than earlier measurements of these fields. There has not yet been any success in interpreting these phenomena in terms of fundamental processes.

A Carrier Telegraph System for Short-Haul Applications. J. L. HYSKO[1], W. T. REA[1] and L. C. ROBERTS[1]. *Elec. Eng.*, **71,** pp. 625–630, July, 1952. (Monograph 2006).

This compact frequency-shift carrier telegraph system provides channels in and above the voice range. The channel terminal unit incorporates arrangements for handling Teletypewriter Exchange Service supervisory signals and employs no electromagnetic relays.

Some Problems in Sampling Accounting Procedure. H. L. JONES[1]. pp. 209–250. *Am. Soc. for Quality Control.* Quality control conference papers. 6th Annual Convention. N. Y., Am. Soc. Quality Control, 1952.

Photometric Determination of Beryllium in Beryllium-Copper Alloys. C. L. LUKE[1] and M. E. CAMPBELL[1]. *Anal. Chem.*, **24,** pp. 1056–1057, June, 1952. (Monograph 2013).

Steady Rotational Flow of Ideal Gases. R. C. PRIM[1]. *Jl. Rational Mech. and Analysis*, **1,** pp. 425–497, July, 1952.

This paper concerns the steady rotational flow of non-viscous and thermally non-conducting gases subject to no extraneous force field. For the most part attention is restricted to gases having constant specific heats. However, some of the results are valid for more general classes of fluids. Uniformity of total flow energy (stagnation enthalpy) or of entropy throughout the flow is not assumed. The present work is intended to be a comprehensive treatment of the status (in 1949) or rotational flow theory from the point of view of the establishment of general properties of such flows and the discovery and study of families of exact solutions to the equations governing them.

Finishing Metal Parts for Telephones. F. B. RINCK[2]. *Metal Progress*, **61,** pp. 65–70, June, 1952.

[1] Bell Telephone Laboratories
[2] Western Electric Company

Binary Counter Uses Two Transistors. R. L. TRENT[1]. *Electronics*, **25,** pp. 100–101, July, 1952.

Various timing and registry functions are provided by transistorized counter with repetition rate from 0 to 50 kc. It has stability without the usual sacrifice in sensitivity and it permits either positive or negative triggering pulses to be used.

Structural Imperfections in Quartz Crystals. W. L. BOND[1] and J. ANDRUS[1]. *Am. Mineral.*, **37,** pp. 622–632, July–August, 1952. (Monograph 2004).

A method for examining the topography of atomic planes is developed and applied to quartz crystals. It is thought to have higher resolution than the method of Wooster and Wooster (*Nature*, **155,** p. 786 (1945)), or that of Ramachandran (*Proc. Ind. Acad. Sci.*, **19A,** p. 280 (1944)). Because of the higher resolution it gives more detailed information. A fair percentage of ostensibly perfect quartz is shown to have slight irregularities.

Packaging Principles Employing Plastics and Printed Wiring to Improve Reliability. W. J. CLARKE[1] and N. J. EICH[1]. pp. 133–137. A. I. E. E., I. R. E. and R. T. M. A. Symposium, Progress in Quality Electronic Components. Proceedings, Wash., D. C., May 5–7, 1952. Wash., D. C., R. T. M. A., 1952.

Miniaturized Components for Transistor Action. P. S. DARNELL[1]. pp. 51–57. A. I. E. E., I. R. E. and R. T. M. A. Symposium, Progress in Quality Electronic Components. Proceedings, Wash. D. C., May 5–7, 1952. Wash., D. C., R. T. M. A., 1952.

Some Basic Concepts of Quality Control. G. D. EDWARDS[1]. Shewhart Medalist Address. *Ind. Quality Control*, **9,** pp. 9–10, July, 1952.

Effective Sum of Multiple Echoes in Television. A. D. FOWLER[1] and H. N. CHRISTOPHER[1]. *S. M. P. T. E., Jl.*, **58,** pp. 491–500, June, 1952.

Observers compared the interfering effect of multiple echoes with that of single echoes in black-and-white television pictures. The multiple echoes were 2, 4 or 8 echoes of equal strength but different delays. The single echoes were 40, 35 or 30 db weaker than the main signal. A method for estimating addition effects of several echoes is presented and demonstrated to be consistent with the test results.

Design Factors Influencing the Reliability of Relays. J. R. FRY[1]. pp. 101–107. A. I. E. E., I. R. E. and R. T. M. A. Symposium, Progress in

[1] Bell Telephone Laboratories

Quality Electronic Components. Proceedings, Wash., D. C., May 5–7, 1952. Wash., D. C., R. T. M. A., 1952.

Energy of a Bloch Wall on the Band Picture. II. Perturbation Approach. C. HERRING[1]. *Phys. Rev.*, **87,** pp. 60–70, July 1, 1952.

The "exchange stiffness" constant, which appears in the theory of the Bloch interdomain wall in ferromagnetics, can be calculated by computing the response of a saturated specimen to a small spatially varying perturbing field. This calculation is carried out here in the self-consistent field approximation, using running waves for the one-electron states, and the result is interpreted physically in terms of precession of the spins of moving electrons. Combination of the present theory with the Stoner-Wohlfarth model of the ferromagnetic electrons in nickel does not give satisfactory results, probably because the latter model does not approximate the actual self-consistent field solution very well. However, application of the theory to the free electron gas is of interest as a confirmation of the validity of the perturbation approach. It is shown that there exist, even in a ferromagnetic metal, quantum states orthogonal to all the low-lying states of the conventional band picture and having the properties of spin waves. The presumably universal relation between the exchange stiffness constant and the energies of spin waves of long wavelength is verified in the present approximation. It is shown that spin waves carry a current in a metal, though not in an insulator. For spin waves of long wavelength the present theory can be shown to include Slater's theory of spin waves in a ferromagnetic insulator, and a fortiori to include all previous theories based on the atomic model.

Nonsynchronous Pulse Multiplex System. A. L. HOPPER[1]. *Electronics*, **25,** pp. 116–120, August, 1952.

Voice transmitters use one frequency simultaneously but no synchronizing pulse is necessary, although time-division multiplexing is used. Random samples from each transmitter are tagged for identification at proper receiver. System is applicable to rural telephony and moving-vehicle communication.

Design of Modulation Equipment for Modern Single-Sideband Transmitters. A. E. KERWIEN[1]. *I. R. E., Proc.*, **40,** pp. 797–803, July, 1952. (Monograph 2012).

This paper deals with considerations that go into the design of modulation equipment for a single-sideband radiotelephone transmitter in which filters are used for sideband suppression. Balance requirements, frequency stability, the choice of intermediate frequencies, and methods of avoiding transmission of spurious frequencies are among the factors which are discussed.

A Multichannel Single-Sideband Radio Transmitter. L. M. KLENK[1], A. J. MUNN[1], and J. NEDELKA[1]. *I. R. E. Proc.*, **40,** pp. 797–803, July, 1952. (Monograph 2012).

This paper describes a new single-sideband radio transmitter for transoceanic

[1] Bell Telephone Laboratories

service which represents a substantial improvement over past design. Its important features include: (a) a frequency band which permits deriving four telephone channels, if desired; (b) a push-button method for changing frequencies within a matter of seconds; (c) an increase in power over its predecessor; and (d) all-around improved transmission performance.

Photometric Determination of Aluminum in Lead, Antimony, and Tin and Their Alloys. C. L. LUKE[1]. *Anal. Chem.*, **24,** pp. 1122–1126, July, 1952. (Monograph 2013).

The work was undertaken because of a need for a reliable method for the determination of traces of aluminum in lead, antimony, and tin and their alloys. As a preparatory step toward the development of such a method, a thorough study of the specificity of the aluminon-thioglycolic acid and the oxine-cyanide-peroxide photometric aluminum methods was made. As a result, an accurate specific method for aluminum has been developed. This method is applicable to the analysis of lead, antimony, and tin and their alloys and can also be adapted for use in the analysis of a wide variety of other ferrous and nonferrous alloys.

Photometric Determination in Manganese Bronze, Zinc Die Casting Alloys, and Magnesium Alloys. C. L. LUKE[1]. and K. C. BRAUN[4]. *Anal Chem.* **24,** pp. 1120–1122, July, 1952. (Monograph 2013).

The work was undertaken in an effort to produce a rapid reliable method for the determination of aluminum appearing as a major constituent in copper, zinc, and magnesium alloys. The work shows that the photometric aluminum method described by Craft and Makepeace is very satisfactory and that by employing thioglycolic acid as a complexing agent it is possible to simplify the usual photometric methods for the determinatior of aluminum in nonferrous alloys. The paper contains experimental material that will aid future workers in the application of this method to other materials.

Amplifiers for Multichannel Single-Sideband Radio Transmitters. N LUND[1], C. F. P. ROSE[1] and L. G. YOUNG[1]. *I. R. E., Proc.*, **40,** pp. 790–796, July, 1952. (Monograph 2012).

Considerations are given for designing high-frequency amplifiers whose performance will meet the high standards required for amplifying multichannel signals. A relationship between tone and speech data is presented to show how the tone rating of the amplifier can be determined from the speech rating and interchannel modulation noise requirements.

Measurement of Dynamic Shear Viscosity and Stiffness of Viscous Liquids by Means of Traveling Torsional Waves. H. J. MCSKIMIN[1]. *Acoustical Soc. Am. Jl.*, **24,** pp. 355–365, July, 1952.

A short periodically repeated train of torsional waves is transmitted along a glass or metal cylindrical rod. After reflection from the free end, these waves are

[1] Bell Telephone Laboratories
[4] American Smelting and Refining Company, Barber, N. J.

sent back to the quartz crystal which serves as both transmitter and receiver. The phase shift and added attenuation caused by immersing the rod in the test liquid are measured by means of a special balancing arrangement, and yield a calculation of the impedance presented to the rod surface. From an analysis of wave propagation both in the rod and in the liquid, one can calculate the characteristic shear impedance of the liquid, and the dynamic viscosity and stiffness. Data for polyisobutylene liquids with static viscosities up to 2000 poises are given for the frequency range 25–150 kc. High frequency data (5–25 mc) for the same liquids obtained by a method previously reported on (see reference 10 (b)) are correlated to the present work. Some results for polypropylene, polyisoprene, polybutadiene, and polypropylene sebacate are also given.

New Transistors Give Improved Performance. J. A. MORTON[1]. *Electronics*, **25,** pp. 100–103, August, 1952.

Better manufacturing processes and germanium materials have provided greater reliability and reproducibility and improved frequency response. Higher power output and better noise figure for high-sensitivity applications are properties of new types.

Microwaves. J. R. PIERCE[1]. *Sci. Am.*, **187,** pp. 43–51, August, 1952.

They are radio waves that range in length from about a quarter of an inch to two feet. Investigated during the war for their utility in radar, they are now widely applied in communication.

Glass Unit for Liquid and Vapor Phase Extraction Employing a Single Processing Chamber. H. A. SAUER[1]. *Anal. Chem.*, **24,** p. 1232, July, 1952.

The Transistors Development Status at Bell Telephone Laboratories, with Demonstration. W. R. SITTNER[1]. pp. 138–142. A. I. E. E., I. R. E. and R. T. M. A. Symposium, Progress in Quality Electron Components. *Proceedings*, Wash., D. C., May 5–7, 1952. Wash., D. C., R. T. M. A., 1952.

Polyethylene Terephthalate as a Capacitor Dielectric. M. C. WOOLEY[1], G. T. KOHMAN[1] and W. MCMAHON[1]. *Elec. Eng.*, **71,** pp. 715–717, Aug., 1952.

Polyethylene terephthalate, or "Mylar", is a new rival of paper for use as the dielectric in electric capacitors. It appears superior in regard to insulation resistance, temperature coefficient of capacitance, and operating temperature range.

[1] Bell Telephone Laboratories

Contributors to this Issue

A. Eugene Anderson, B.Sc. in E.E., Ohio State University, 1939; M.Sc., Ohio State University, 1939. U. S. Army, 1942–46. Bell Telephone Laboratories, 1939–. Mr. Anderson is concerned with the development of semi-conductor devices, including the transistor. In the past he has been engaged in microwave electron tube and electron beam tube development. Member of I. R. E., Tau Beta Pi, Sigma Xi, Eta Kappa Nu, and Sigma Pi Sigma.

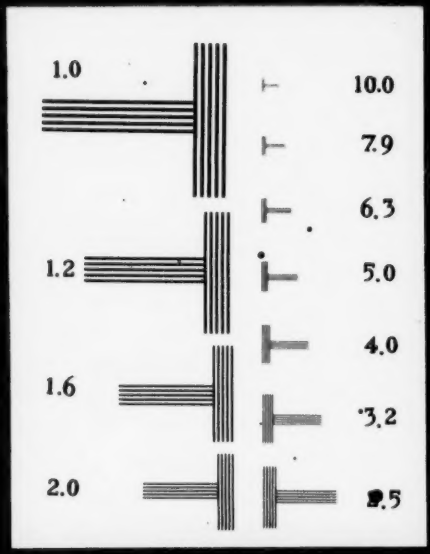
Arthur C. Keller. B.S., Cooper Union, 1923; M.S., Yale University, 1925; E.E., Cooper Union, 1926; Columbia University, 1926–30; Western Electric Company, 1917–25; Bell Telephone Laboratories, 1925–. Special Apparatus Development Engineer, 1943; Switching Apparatus Development Engineer, 1946; Assistant Director of Switching Apparatus Development, 1949; Director of Switching Apparatus Development, 1949–. Mr. Keller's experience in the Bell System includes development and design of telephone instruments; development of systems and apparatus for recording and reproducing sound; and, during World War II, the development, design, and preparation for manufacture of sonar systems and apparatus. His department, in addition to being responsible for a number of military projects, is responsible for the fundamental studies of switching apparatus and the development, design, and preparation for manufacture of electromagnetic and electromechanical switching apparatus for telephone systems. Member of the American Physical Society, A. I. E. E., Acoustical Society of America, I. R. E., S. M. P. T. E., and the Yale Engineering Association. Representative for Bell Telephone Laboratories in the Society for Experimental Stress Analysis. For his contributions to the Navy during World War II, he received awards from the Bureau of Ships and the Bureau of Ordnance.

Samuel P. Morgan, Jr., B.S., California Institute of Technology, 1943; M.S., California Institute of Technology, 1944; Ph.D., California Institute of Technology, 1947. Bell Telephone Laboratories, 1947–. A research mathematician, Dr. Morgan specializes in electromagnetic theory. He has been particularly concerned with problems of wave guide and coaxial cable transmission. Member of the American Physical Society, Tau Beta Pi, and an associate member of Sigma Xi.

OSCAR MYERS, B. Chem., Cornell University, 1921. Western Electric Company, 1921–24. Bell Telephone Laboratories, 1924–. Mr. Myers' early work was concerned with circuit testing. He then worked in a circuit design group, from 1929 until 1948. Since 1948 he has been a member of the Switching Engineering Department, and has contributed to the design or development of practically all of the switching developments of the Laboratories, particularly in the field of common controls. His work has covered panel, crossbar, automatic message accounting, toll, crossbar tandem, and other systems. Member of A. I. E. E.

W. RAE YOUNG, JR., B.S., in E.E., University of Michigan, 1937; Bell Telephone Laboratories, 1937–. Mr. Young is in the Systems Studies Department, where he is giving consideration to new system possibilities for meeting future communication needs. During World War II, he worked in radar development and, later, on systems problems in radio communications. From 1945–50 Mr. Young helped set up Bell System performance requirements for mobile radio telephone equipment. Member of I. R. E. and Sigma Xi.

RESOLUTION CHART



100 MILLIMETERS

INSTRUCTIONS Resolution is expressed in terms of the lines per millimeter recorded by a particular film under specified conditions. Numerals in chart indicate the number of lines per millimeter in adjacent "T-shaped" groupings.

In microfilming, it is necessary to determine the reduction ratio and multiply the number of lines in the chart by this value to find the number of lines recorded by the film. As an aid in determining the reduction ratio, the line above is 100 millimeters in length. Measuring this line in the film image and dividing the length into 100 gives the reduction ratio. Example: the line is 20 mm. long in the film image, and $100/20 = 5$.

Examine "T-shaped" line groupings in the film with microscope, and note the number adjacent to finest lines recorded sharply and distinctly. Multiply this number by the reduction factor to obtain resolving power in lines per millimeter. Example: 7.9 group of lines is clearly recorded while lines in the 10.0 group are not distinctly separated. Reduction ratio is 5, and $7.9 \times 5 = 39.5$ lines per millimeter recorded satisfactorily. $10.0 \times 5 = 50$ lines per millimeter which are not recorded satisfactorily. Under the particular conditions, maximum resolution is between 39.5 and 50 lines per millimeter.

Resolution, as measured on the film, is a test of the entire photographic system, including lens, exposure, processing, and other factors. These rarely utilize maximum resolution of the film. Vibrations during exposure, lack of critical focus, and exposures yielding very dense negatives are to be avoided.